

ISSN 1854-6250

APEM
journal

Advances in Production Engineering & Management

Volume 15 | Number 3 | September 2020





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
Qingliang Meng, Jiangsu University of Science and Technology, China
Lanndon A. Ocampo, Cebu Technological University, Philippines
Iztok Palcic, University of Maribor, Slovenia
Krsto 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

Limited Permission to Photocopy: Permission is granted to photocopy portions of this publication for personal use and for the use of clients and students as allowed by national copyright laws. This permission does not extend to other types of reproduction nor to copying for incorporation into commercial advertising or any other profit-making purpose.

Subscription Rate: 120 EUR for 4 issues (worldwide postage included); 30 EUR for single copies (plus 10 EUR for postage); for details about payment please contact: info@apem-journal.org

Cover and interior design: Miran Brezocnik

Printed: Tiskarna Koštomaj, Celje, Slovenia

Subsidizer: The journal is subsidized by Slovenian Research Agency

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.

Copyright © 2020 CPE, University of Maribor. All rights reserved.

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



University of Maribor
Chair of Production Engineering (CPE)

Advances in Production Engineering & Management

Volume 15 | Number 3 | September 2020 | pp 251–372

Contents

Scope and topics	254
An exact approach for the consistent vehicle routing problem (ConVRP)	255
Barros, L.; Linfati, R.; Escobar, J.W.	
Decentralized optimization of the flexible production lines	267
Malega, P.; Rudy, V.; Kanász, R.; Gazda, V.	
Interactive impacts of overconfidence and fairness concern on supply chain performance	277
Zhang, Z.J.; Wang, P.; Wan, M.Y.; Guo, J.H.; Luo, C.L.	
Computational analysis of cavitation at the tongue of the volute of a centrifugal pump at overload conditions	295
Hu, Q.; Yang, Y.; Cao, W.	
Optimization of process performance by multiple pentagon fuzzy responses: Case studies of wire-electrical discharge machining and sputtering process	307
Al-Refaie, A.; Lepkova, N.; Abbasi, G.; Bani Domi, G.	
An integrated CNC system for chatter suppression in turning	318
Jasiewicz, M.; Miądlicki, K.	
Simulation-based time evaluation of basic manual assembly tasks	331
Turk, M.; Pipan, M.; Šimic, M.; Herakovič, N.	
Manufacturer's customer satisfaction incentive plan for duopoly retailers with Cournot or collusion games	345
Hu, H.; Zhang, Z.; Wu, Q.; Han, S.	
Testing of novel nano gold ink for inkjet printing	358
Rudolf, R.; Majerič, P.; Golub, D.; Tiyyagura, H.R.	
Calendar of events	369
Notes for contributors	371

Journal homepage: apem-journal.org

ISSN 1854-6250 (print)

ISSN 1855-6531 (on-line)

©2020 CPE, University of Maribor. All rights reserved.

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

An exact approach for the consistent vehicle routing problem (ConVRP)

Barros, L.^a, Linfati, R.^b, Escobar, J.W.^{c,*}

^aEscuela de Ingeniería Industrial, Universidad del Bío-Bío, Chile, Concepción, Chile

^bDepartamento de Ingeniería Industrial, Universidad del Bío-Bío, Chile, Concepción, Chile

^cDepartment of Accounting and Finance, Faculty of Business Management, Universidad del Valle, Cali, Colombia

ABSTRACT

This paper proposes a mathematical model for the Consistent Vehicle Routing Problem (ConVRP). The ConVRP is an extension of the VRP, considering customer satisfaction through consistent service. The consistency may be based on time or on the vehicle that offers the service. This paper proposes a novel mathematical model that allows solving the ConVRP for several companies for which visits to the customers need to be from the same service provider (namely, the same vehicle and driver). The efficiency of the model is tested on structured instances by changing customer distribution (uniform or clustered), depot location, and arrival time to the customer and removing certain constraints to see if they affect the performance of the objective function. The mathematical model is flexible and could be adapted to any characteristic of instances. The model was developed in the AMPL programming language and solved with the solvers CPLEX and Gurobi. The results are promising based on the efficiency of the proposed method at solving the problem.

© 2020 CPE, University of Maribor. All rights reserved.

ARTICLE INFO

Keywords:

Vehicle routing problem (VRP);
Consistent vehicle routing (ConVRP);
Mathematical model;
Mixed integer linear programming model;
Optimization;
Exact algorithms;
Modelling;
CPLEX;
Gurobi

*Corresponding author:

john.wilmer.escobar@correounivalle.edu.co
(Escobar, J.W.)

Article history:

Received 24 March 2020

Revised 10 October 2020

Accepted 13 October 2020

1. Introduction

The main objective of the consistent vehicle routing problem (ConVRP) is customer satisfaction through consistent service. The consistency may be based on time or on the vehicle that offers the service. There is extensive literature on methods to solve the well-known vehicle routing problem (VRP) and its variants [1], but few studies are related to ConVRP. The VRP is critical in successful logistics execution. The emergence of technologies and information systems, allowing for seamless mobile and wireless connectivity between delivery vehicles and distribution facilities, has generated a new research field integrating real-time VR with consistent distribution management [2].

The study of logistical problems such as vehicle routing and its different approaches is gaining more interest both in academia and its application to real issues [3, 4]. Although cost minimization is traditionally only considered in a VRP problem, other objectives can be defined for the same problem (distance, time, capacity, etc.). There are also other decision-making criteria of a more significant impact, such as delays or delivery schedules. Gaining a better market position requires focusing more on the ConVRP, whose objective is to emphasize customers through a more consistent service [5].

This paper proposes a mathematical mixed-integer linear programming model to solve the ConVRP. The proposed model can determine the total travel time of a vehicle fleet for a certain number of specific days, which routes should be taken by each vehicle per day, and the vehicles' arrival time for each customer. The mathematical model was tested on structured instances by analyzing changes to aspects such as customer distribution (uniform or clustered), depot location, arrival time at the customer, and removing certain constraints that affect the performance of the objective function. The model was developed in the AMPL programming language and solved with the solvers CPLEX and Gurobi. Although a mathematical model for the ConVRP attracts growing attention due to the few studies related to the problem. The mathematical structure of the model is novel, and to the best of the author's knowledge, this is the first time a model for the ConVRP solves real and structured instances for companies that provide a consistent service over time, such as courier companies, elderly care service companies, and cleaning sectors, under realistic constraints.

The paper is organized as follows: Section 2 gives an overview of the literature on the problem of ConVRP. Section 3 presents the proposed mathematical model with its considerations and assumptions, and Section 4 presents the computational results obtained from instances. Finally, the conclusions and future work are addressed in Section 5.

2. Literature review

This study first investigates the ConVRP in known instances from the literature, through an analysis of the VRP, to apply the method subsequently to real problems that are similar to the issue analyzed. The VRP consists of determining a set of routes for a vehicle fleet that leaves one or more depots to meet the demand of various geographically dispersed customers [6–8]. Usually, the VRP consists of determining the best routes and/or assignments to deliver goods or services to geographically distributed customers. This problem involves assigning a group of customers to a group of vehicles and drivers and sequencing their visits. The objective of the VRP is to deliver a product or service with the minimum total cost involved in the routes [9].

According to [10], there are other important components when defining a route problem, such as the type of network and objectives to optimize. Concerning the kind of structure, problems are classified according to node covering or arc covering issues. It is also possible for the objective to be to minimize total cost, total time, or total distance traveled and, at the same time, maximize the quality of service or the profits from charging customers.

There are different variants of the VRP: Multi-Depot VRP [8], Multi-Depot VRP with Heterogeneous Fleet [11], Periodic VRP [12], Split-Delivery VRP [13], Split-Delivery VRP with Time Windows [14], Heterogeneous VRP with Time Windows [15], stochastic VRP [16], VRP with Pickup and Delivery [17], VRP with Backhauls [18], VRP with time windows [19], and ConVRP [20]. ConVRP is defined as an extension of periodic VRP in which minimum-cost routes are designed to deliver a service to a group of customers with known demand over time through a homogeneous fleet of qualified vehicles. Customers may receive the service either once or at a pre-defined frequency; however, frequent customers must receive consistent service over a planning period, such that the same vehicle consistently satisfies the time windows and service [5, 21–22].

According to [21], in the ConVRP, each customer must receive the service on a specific day of a set D . Each customer can receive one service no more than once a day from any of the K identical vehicles. When a customer receives the service, the same driver will make the delivery at roughly the same time during the planning horizon of D days, so the maximum variation of the arrival time is no more than L time units. Each vehicle has a capacity of Q units and cannot operate for more than T units of time. The objective is to develop a set of routes for the fleet that minimizes the total operating time of the vehicle over D days.

The study by [23] demonstrated that some of the solutions provided by Groër et al. (2009) are not feasible and proposed an algorithm called ConRTR that obtained viable solutions. The results showed that with a relatively small increase in total travel time, the solutions provided by ConRTR were much more consistent than those proposed by [22].

ConVRP has numerous real-life applications in which visits to the customers need to be from the same service provider (namely, the same vehicle and driver). Also, in many cases, customers need to be visited at a given frequency. Some real-life examples of this type of service factor are parcel delivery and courier services, elderly care services, and cleaning services. In these cases, the primary endeavor is to gain a competitive advantage by bonding with the customer. In [3], it is considered the problem of consistency in the transport service for persons with disabilities. The authors presented a VRP that was consistent over time and solved it with a neighborhood search heuristic. Each iteration of the heuristic required solving a complex VRP with multiple time windows and no wait time.

The research in [24] pointed out that one of the most common problems at present is the transportation of persons with mental disabilities to rehabilitation centers. Because of the patients' lack of autonomy, most of those affected do not have a vehicle and cannot use public transport. Consequently, social health centers generally use specialized transport companies to organize and carry out daily trips. Customers need to be transported regularly, but sometimes with some variations during the week. These users are particularly sensitive to changes, so it is essential to keep consistent schedules and to provide the service by the same driver to create a bond of trust and gain a competitive advantage.

The paper [25] proposed a heuristic for a ConVRP in a distribution center of a food company. The problem was characterized by a group of customers that varied from day to day, and by their demand. Besides, each customer requires orders delivered within a given time period.

In [26], it is developed as a mathematical model to address the ConVRP of multiple daily deliveries and different service-level configurations, such as time windows and release dates. A mathematical programming-based decomposition approach was proposed to solve the problem. The algorithm was tested with real data from a Portuguese pharmaceutical distribution company. [27] proposed a metaheuristic algorithm based on local search for scheduling and routing of mobile nurses. Mobile nurses visited patients regularly at home. Some patients require up to three visits per day on predefined time windows. The visits required different qualifications, and the nurses had different qualification levels. This problem involved consistency in the visits.

In [28], an exact algorithm based on column generation was proposed for the ConVRP. Each variable of this model represented the set of routes assigned to a vehicle over the planning horizon. The proposed algorithm could solve small and medium-sized instances with five planning periods and 30 customers.

The work [5] presents a solution framework for the ConVRP. In this problem, customers received service once or at a predefined frequency, but frequent customers had to receive consistent service; that is, they had to be visited by the same driver throughout the planning period. The proposed algorithm adopted a two-level decomposition scheme. The computational experiments with benchmark data sets illustrated the competitiveness of the proposed approach compared to current results.

The paper [29] proposes a variable neighborhood search (VNS) for the ConVRP. The proposed algorithm consisted of two stages. In the first stage, VNS was applied to obtain approximate solutions. The obtained solutions might be unfeasible; if a solution had acceptable quality, the second step attempted to improve it. The proposed algorithm was tested on the benchmark ConVRP data set and compared with existing ConVRP approaches from the literature. The results showed that VNS outperformed all existing ConVRP methods regarding the quality of the solutions obtained.

The work proposed by [30] studied a variant of the traditional ConVRP considering profits. There were two sets of customers in this problem, the frequent customers that must be serviced and the infrequent potential customers with known and estimated profits, both having known demands and service requirements over a planning horizon of multiple days. The objective was to determine the vehicle routes that maximized net profit while satisfying vehicle capacity, route duration, and consistency constraints. This paper proposed a new mathematical model that captures the profit-collecting nature and other features of the problem and an adaptive tabu search that used short- and long-term memory structures to guide the search process. In [31], the generalized ConVRP (GenConVRP) is considered, where a limited number of drivers visited each

customer, and variation in arrival times was penalized in the objective function. The vehicle departure times could be adjusted to obtain stable arrival times.

Additionally, customers were associated with time windows. Unlike the previous work on the ConVRP, this study used the template concept to generate route plans. The proposed approach was based on a flexible neighborhood search. Several destroy-and-repair heuristics were designed to remove customers from the routes and to reinsert them in better positions. The multi-purpose version of the GenConVRP is reviewed in [32].

3. Problem description and proposed approach

3.1 ConVRP description

Most studies in the ConVRP literature emphasize many logistics companies' current problems: low customer satisfaction with the company's performance stemming from its sole focus on total cost minimization and not caring about customer service. For this reason, companies now seek to increase service quality even if it means an increase in costs.

The ConVRP has great relevance and importance to that problem. The main objective of the ConVRP is the minimization of the travel time of vehicles on the days required. In a general context, the goal is to meet customers' individual demands in the defined planning horizon, leaving the customer satisfied and achieving consistent times and drivers. As mentioned above, it is essential that each customer receive service by the same driver each time, since this creates a bond of trust with the customer and improves customer satisfaction with the service offered. Moreover, there is a time constraint on each vehicle's arrival to the customer, since the customer should receive service simultaneously on each service day to achieve consistency concerning time [21]. The critical elements of the ConVRP are:

- Unequal hours of service delivery to the customer.
- Undefined routes for each vehicle.
- Inconsistency in the drivers who deliver service to the customers.

3.2 Proposed mathematical model

The proposed model solves real and structured instances for companies that provide a consistent service over time, such as companies in the courier, elderly care, and cleaning sectors, under realistic constraints. The proposed mixed-integer linear programming model determines customers' routes to visit, the vehicles assigned, and the arrival times to the customers.

Sets

D	Planning horizon, equivalent to the set of days the service is required.
V	Available vehicles (homogeneous fleet)
$Nodes$	Set of nodes (depot plus customers)
$Arcs$	Set of arcs between nodes (physical connections)

Parameters

T	Daily capacity in hours per vehicle
S_{id}	Service time for customer i on day d
L	Maximum arrival-time variation between two customers
Q	Maximum capacity of each vehicle
q_{id}	Demand of customer i on day d
w_{id}	1 indicates customer i requires service on day d , 0 indicates otherwise
t_{ij}	Travel time associated with arc (i, j)

Decision variables

$$x_{ijkd} = \begin{cases} 1 & \text{if vehicle } k \text{ visits customer } j \text{ after customer } i \text{ on day } d \\ 0 & \text{otherwise} \end{cases}$$

$$y_{ikd} = \begin{cases} 1 & \text{if customer } i \text{ is visited by vehicle } k, \text{ on day } d \\ 0 & \text{otherwise} \end{cases}$$

a_{id} is arrival time to customer i on day d ($i = 0$ at the depot)

u_{id} is auxiliary variable for subtour elimination

Objective function

$$\text{Minimize } \sum_{d \in D} \sum_{k \in V} \sum_{(i,j) \in \text{Arcs}} t_{ij} x_{ijkd} \tag{1}$$

Eq. 1 seeks to minimize the total travel time of the vehicles to the customers over the whole planning horizon.

Constraints

$$y_{0kd} = 1 \quad \forall k \in V \forall d \in D \tag{2}$$

$$a_{0d} = 0 \quad \forall d \in D \tag{3}$$

$$\sum_{k \in V} y_{ikd} = w_{id} \quad \forall i \in \text{Nodes} \forall d \in D, i \geq 1 \tag{4}$$

$$\sum_{i \in \text{Nodes}} q_{id} * y_{ikd} \leq Q \quad \forall i \in \text{Nodes} \forall d \in D \forall k \in V, \quad i \geq 1 \tag{5}$$

$$\sum_{i \in \text{Nodes}} x_{ijkd} = y_{jkd} \quad \forall j \in \text{Nodes} \forall k \in V \forall d \in D, \quad i \geq 1 \tag{6}$$

$$\sum_{i \in \text{Nodes}} x_{jikd} = y_{jkd} \quad \forall j \in \text{Nodes} \forall k \in V \forall d \in D, \quad i \geq 1 \tag{7}$$

$$w_{id_\alpha} + w_{id_\beta} - 2 = y_{ikd_\alpha} - y_{ikd_\beta} \quad \forall d_\alpha, d_\beta \in D \forall i \in \text{Nodes} \forall k \in V, \alpha \neq \beta \tag{8}$$

$$w_{id_\alpha} + w_{id_\beta} - 2 = y_{ikd_\beta} - y_{ikd_\alpha} \quad \forall d_\alpha, d_\beta \in D \forall i \in \text{Nodes} \forall k \in V, \alpha \neq \beta \tag{9}$$

$$a_{id} + x_{ijkd}(s_{id} + t_{ij}) - (1 - x_{ijkd}) * T \leq a_{jd} \quad \forall i, j \in \text{Arcs} \quad k \in V \quad d \in D, \quad j \geq 1 \tag{10}$$

$$a_{id} + x_{ijkd}(s_{id} + t_{ij}) - (1 - x_{ijkd}) * T \geq a_{jd} \quad \forall i, j \in \text{Arcs} \quad k \in V \quad d \in D, \quad j \geq 1 \tag{11}$$

$$a_{id} + w_{id} * (s_{id} + t_{i0}) \geq 0 \quad \forall i \in \text{Nodes} \forall d \in D, i \geq 1 \tag{12}$$

$$a_{id} + w_{id} * (s_{id} + t_{i0}) \leq T * w_{id} \quad \forall i \in \text{Nodes} \forall d \in D, i \geq 1 \tag{13}$$

$$a_{id_\alpha} - a_{id_\beta} \leq L + T * (2 - w_{id_\alpha} - w_{id_\beta}) \quad \forall d_\alpha, d_\beta \in D \forall i \in \text{Nodes}, \alpha \neq \beta \tag{14}$$

$$a_{id_\beta} - a_{id_\alpha} \leq L + T * (2 - w_{id_\alpha} - w_{id_\beta}) \quad \forall d_\alpha, d_\beta \in D \forall i \in \text{Nodes}, \alpha \neq \beta \tag{15}$$

$$u_{id} + 1 \leq u_{jd} + \text{Nodes} * (1 - x_{ijkd}) \quad \forall i, j \in \text{Arcs} \quad k \in V \quad d \in D, \quad j \geq 1 \tag{16}$$

$$x_{ijkd} \in \{0,1\} \quad \forall i, j \in \text{Nodes} \forall k \in V \forall d \in D \tag{17}$$

$$y_{ikd} \in \{0,1\} \quad \forall i \in \text{Nodes} \forall k \in V \forall d \in D \tag{18}$$

$$a_{id} \geq 0 \quad \forall i \in \text{Nodes} \forall d \in D \tag{19}$$

$$u_{id} \geq 0 \quad \forall i \in \text{Nodes} \forall d \in D \tag{20}$$

Eq. 2 determines that each vehicle must visit the depot first and then go to its respective customer. Eq. 3 corresponds to the arrival time of the vehicle at the depot, which must be equal to 0. This ensures that the vehicles visit the depot at time 0 every day. Eq. 4 ensures that customers

are visited exactly once when the service is required, that is, each customer is visited once per service day. Constraints 5 limit the capacity of the vehicles to prevent them from exceeding the Q units of product the vehicles can deliver per day. Eqs. 6 and 7 ensure that each customer has only one predecessor and successor.

The following constraints achieve model consistency with the customers that will be visited in the planning horizon. Eqs. 8 and 9 ensure that each customer receives the service by the same driver each time the customer requests service. Constraints 10 and 11 determine the daily arrival times of individual customers, which is necessary to achieve consistency in the time windows. Eqs. 12 and 13 limit the time the vehicles circulate, which translates to the maximum travel time of each vehicle. Constraints 14 and 15 are those that achieve time consistency, as they indicate that the difference in arrival times to customer i between days α and β cannot be more than L time units. This ensures customers are serviced on both days at roughly the same time. The value of L is initially adjusted to be the same as T , which leads to a relaxation of the problem, since there is no consistency in the times but there is in the vehicles. Therefore, when L has that value, it indicates that the vehicle can deliver the service within the daily limit of hours in circulation. Later, the L value is adjusted to a number lower than T to achieve consistency in the service delivery time. Constraint 10 seeks to eliminate subtours with Eq. 16, which is adapted from [33]. Computational tests were conducted on removing Constraints 16 and checking the behavior of the model. Constraint 11 keeps vehicles constantly moving to meet arrival times. Finally, Eqs. 17 and 18 output binary variables, while Eqs. 19 and 20 output nonnegative variables, which means that the variables cannot be negative.

4. Results, analysis and discussion

The following section shows the different results obtained from different structural instances under the conditions explained below. Experiments were performed with test instances to verify the proper operation of the proposed algorithm. The general computational results are analyzed for different structural instances. Different solver options are also tested to analyze other results based on computation times or objectives. The proposed model was developed in AMPL language version 20180822 and solved with the solvers CPLEX 12.8.0 and Gurobi 8.0.0 in an Intel Xeon E5-2660 v2 Dual Core @ 2.20 GHz computer with 64 GB RAM and a 64-bit Red Hat Enterprise Linux 7 operating system.

4.1 Instances definition

The instances are developed taking into account the following considerations:

- Customers: A total of three groups, of 10, 15, and 20 customers, are considered, in addition to uniform and cluster distributions.
- Depot: It is located at the center of the data or at one of the ends of the region.
- Arrival time difference (L): Instances are created with different L parameters, which are $L = T$, $L = 3$, and $L = 1$.
- Constraints: A modification is made to the model by removing Constraints 16 and 11.

Certain conditions mentioned by [21] and [34] in the literature are established when the instances are created; those conditions are the following:

- The uniform distributions of customer locations are generated within a square with vertices (0,0), (10,0), (10,10), (0,10). Cluster instances are generated according to the method proposed by [34]. This procedure consists of four steps:
 1. Two clusters are distributed uniformly between the coordinates mentioned above.
 2. Then customers are generated uniformly, and the Euclidean distance between the point and the cluster is obtained using (21):

$$D = \sqrt{(C_x - x)^2 + (C_y - y)^2} \quad (21)$$

3. To verify that the point is close to the cluster, Eq. 22 is used:

$$P_1 = e^{\left\{\frac{-D(p,C_1)}{0,8}\right\}} + e^{\left\{\frac{-D(p,C_2)}{0,8}\right\}} \tag{22}$$

4. Finally, P_1 is compared with $P_2 = U(0,1)$. If $P_1 < P_2$ then the point is added; otherwise, the point is discarded, and the same process is repeated until all customers are captured.

- The planning horizon is set to 3 days.
- The probability of customers requiring service every day is 70 %.
- For customers that require service on a certain day, demand is uniformly distributed in (1,3).
- All service times are set to one unit, which varies depending on the total number of customers.
- Vehicle capacity is homogeneous, and the maximum capacity is 15 units.

A total of 108 instances are generated, divided into four types.

4.2 Structured instances results

All instances are evaluated with a maximum calculation time of 3600 seconds (1 hour) since that time is considered reasonable to find an optimal solution. For the notations on the tables, the term "CPU Time" corresponds to the time in seconds taken to find the result, "Gap" is the percentage difference between the best number obtained and the best integer solution found by the model, and "Z" is the value of the objective function.

Customer distribution

This section compares the results obtained when the model is applied to instances focused on uniform and clustered distribution of the customers for a maximum of 20 customers. Table 1 shows the results of the instances for uniformly distributed customers.

The number of vehicles used in each instance achieves balanced routes for each vehicle since the probability of the daily demand for each customer is 70 %. In particular, route balance is not achieved if the number of vehicles is high. Table 2 shows the results of the model for instances generated in clusters.

The analysis of the results obtained with customers distributed uniformly and in clusters shows that this latter model finds better solutions in less computation time. A clustered distribution of customers results in better model behavior and reduced computation times, which is reflected in a more significant number of optimal solutions. Finally, the Gurobi software delivered better results in computation time and Gap percentage, but these results were not more relevant than those obtained with CPLEX.

Table 1 Performance of the model with uniformly distributed customers

Number of customers	Vehicles used	CPLEX			Gurobi		
		Gap (%)	CPU Time (s)	Z (time)	Gap (%)	CPU Time (s)	Z (time)
10	2	0.00	0.25	136.90	0.00	0.56	136.90
15	2	11.68	3600.00	149.60	4.99	3600.00	149.60
20	3	27.16	3600.00	192.25	15.86	3600.00	192.25

Table 2 Performance of the model with customers distributed in clusters

Number of customers	Vehicles used	CPLEX			Gurobi		
		Gap (%)	CPU Time (s)	Z (time)	Gap (%)	CPU Time (s)	Z (time)
10	2	0.00	143.88	135.08	0.00	368.54	135.08
15	2	0.00	13.77	126.32	0.00	12.48	126.32
20	3	7.44	3600.00	151.71	4.13	3600.00	151.71

Depot location

This section shows the results obtained for different depot locations, considering both customer distributions, uniforms, and clustered. In particular, the depot is located at the corner or the center of the customers. In the corner option, the depot is put on one of the vertices that delimit customers' locations. Table 3 shows the results obtained when the model is applied to the instances.

The model finds an optimal solution for 50 % of the instances studied from the data obtained with the solver, in computation time of fewer than 143 seconds. In contrast, the other 50 % of the instances fail to solve optimally, but feasible solutions are still found, of which two are from customers with uniform distribution and, at the same time, have the highest Gap. The other option puts the depot precisely in the center of the coordinates of the customers. Table 4 shows the results obtained when the model is applied to the generated instances.

As shown in Table 4, customers' travel time with the clustered distribution is considerably lower than with the uniform distribution. The model performs better when the depot is located at the center than at a corner. Indeed, the model found optimal solutions for all instances of this type. Additionally, an analysis of the solver results shows that the computation times of Gurobi are shorter in most cases, though in 4 out of 12 instances, the computation time was shorter with CPLEX, which indicates a lower performance for these instances. For this reason, the arrival time and relaxation of constraints cases were not analyzed with the Gurobi solver.

Table 3 Performance of the model for the depot located on a corner

Customers	Distribution	CPLEX			Gurobi		
		Gap (%)	CPU Time (s)	Z (time)	Gap (%)	CPU Time (s)	Z (time)
10	Uniform	0.00	0.25	136.90	0.00	0.56	136.90
15	Uniform	11.68	3600.00	149.60	4.99	3600.00	149.60
20	Uniform	27.16	3600.00	192.25	15.86	3600.00	192.25
10	Cluster	0.00	143.88	135.08	0.00	368.54	135.08
15	Cluster	0.00	13.77	126.32	0.00	12.48	126.32
20	Cluster	7.44	3600.00	151.71	4.13	3600.00	151.71

Table 4 Performance of the model for a depot located at the center

Customers	Distribution	CPLEX			Gurobi		
		Gap (%)	CPU Time (s)	Z (time)	Gap (%)	CPU Time (s)	Z (time)
10	Uniform	0.00	1.15	81.29	0.00	0.86	81.29
15	Uniform	0.00	5.28	112.02	0.00	2.52	112.02
20	Uniform	0.00	1289.33	129.76	0.00	1971.33	129.76
10	Cluster	0.00	1.99	54.47	0.00	2.60	54.47
15	Cluster	0.00	22.62	64.96	0.00	15.59	64.96
20	Cluster	0.00	141.38	79.68	0.00	266.74	79.68

Arrival time differences (L)

The arrival time difference is the parameter that limits how long a vehicle can take to provide service to a customer on different days. We are interested in reducing that time as much as possible to achieve more consistent arrival times. Table 5 shows the results of applying the model with an L value equal to T , and Table 6 shows the results for $L = 3$, considering the depot located at one end and at the center, and with a clustered distribution of customers. Finally, Table 7 shows the results of the instances with $L = 1$.

Table 5 Performance of the model with $L = T$

Customers	Depot	T	L	CPLEX		
				Gap (%)	CPU Time (s)	Z (time)
10	Corner	30	30	0.00	143.88	135.08
15	Corner	38	38	0.00	13.77	126.32
20	Corner	40	40	7.44	3600.00	151.71
10	Center	30	30	0.00	1.99	54.47
15	Center	38	38	0.00	22.62	64.96
20	Center	40	40	0.00	141.38	79.68

Table 6 Performance of the model with $L = 3$

Customers	Depot	T	L	CPLEX		
				Gap (%)	CPU Time (s)	Z (time)
10	Corner	30	3	0.00	25.75	135.08
15	Corner	38	3	0.00	60.81	126.81
20	Corner	40	3	10.50	3600.00	154.26
10	Center	30	3	0.00	2.68	54.47
15	Center	38	3	0.00	82.52	66.72
20	Center	40	3	11.57	3600.00	82.42

Table 7 Performance of the model with $L = 1$

Customers	Depot	T	L	CPLEX		
				Gap (%)	CPU Time (s)	Z (time)
10	Corner	30	1	Unfeasible	13.87	Unfeasible
15	Corner	38	1	16.57	3600.00	146.47
20	Corner	40	1	18.27	3600.00	165.47
10	Center	30	1	0.00	8.94	61.69
15	Center	38	1	27.24	3600.00	82.56
20	Center	40	1	28.81	3600.00	98.93

The results showed that as the L parameter decreases, that is, as the difference in arrival time approaches 0, the model struggles to find solutions. The travel times obtained with the model increase as the difference in arrival time decreases; this increase in the value of the objective function is most clearly seen when $L = 1$. An optimal solution is not achieved for most of these results, but if the computation time had been increased, it would have been possible to obtain results closer to those obtained in the previous instances. In addition, for those cases in which feasible solutions were obtained, the time the solver took to obtain a solution was no more than roughly 3 minutes.

Relaxation of constraints

In the model developed, there are two constraints that need to be analyzed since these can significantly affect the results that will be obtained. In particular, tests were conducted without Constraints 11 and 16, and the analysis was repeated. In the first test, Constraint 16, whose function is to remove the subtour, was eliminated since Restriction 10 can fulfill the same function. By doing this, the relaxed problem can be solved with the instances consisting of clustered customers, both depot locations, and $L = 3$. Tables 8 and 9 show the results for each model.

In the first comparison made with the removal of Constraint 16, no significant changes are seen in the solver; in fact, the results obtained are very similar, both in Gap value and computation times. Therefore, removing this constraint is no longer considered since it does not significantly improve the original model.

Table 8 Performance of the original model

Customers	Depot	L	CPLEX		
			Gap (%)	CPU Time (s)	Z (time)
10	Corner	3	0.00	25.75	135.08
15	Corner	3	0.00	60.81	126.81
20	Corner	3	10.50	3600.00	154.26
10	Center	3	0.00	2.68	54.47
15	Center	3	0.00	82.52	66.72
20	Center	3	11.57	3600.00	82.42

Table 9 Performance of the model without Constraint 16

Customers	Depot	L	CPLEX		
			Gap (%)	CPU Time (s)	Z (time)
10	Corner	3	0.00	47.09	135.08
15	Corner	3	0.00	44.36	126.81
20	Corner	3	10.23	3600.00	154.26
10	Center	3	0.00	2.98	54.47
15	Center	3	0.00	144.68	66.72
20	Center	3	12.83	3600.00	82.42

The second test was to eliminate constraint 11, which was intended to keep the vehicle in constant motion along the entire route; therefore, if this constraint is eliminated, then the vehicle will wait in its place before the next customer is visited. As a result, the instances of customers organized in clusters, both depot locations, and $L = 1$ were considered. Tables 10 and 11 show the original model results and the model without Constraint 11.

This comparison shows that Constraint 11 can be eliminated if the driver is to wait before traveling to the next customer, to provide consistent arrival times in less travel time. The results are better than those of the original model since the objective function's values are lower, which indicates a lower total travel time and, thus, a lower cost. The results obtained for vehicles' daily travel times show that the model performs better with a clustered distribution of customers and a centrally located depot since total travel times are the lowest out of all experiments. Besides, total travel times rise as the difference in arrival times to customers approaches 0.

Table 10 Performance of the original model

Customers	Depot	L	CPLEX		
			Gap (%)	CPU Time (s)	Z (time)
10	Corner	1	Unfeasible	13.87	Unfeasible
15	Corner	1	16.57	3600.00	146.47
20	Corner	1	18.27	3600.00	165.47
10	Center	1	0.00	8.94	61.69
15	Center	1	27.24	3600.00	82.56
20	Center	1	28.81	3600.00	98.93

Table 11 Performance of the model without Constraint 11

Customers	Depot	L	CPLEX		
			Gap (%)	CPU Time (s)	Z (time)
10	Corner	1	0.00	11.12	135.08
15	Corner	1	0.00	6.61	126.32
20	Corner	1	8.15	3600.00	151.91
10	Center	1	0.00	2.06	54.47
15	Center	1	0.00	29.01	65.30
20	Center	1	0.00	118.48	79.68

5. Conclusion

This article proposes a mathematical model for the ConVRP. The ConVRP is a variant of the capacitated VRP where customers can receive service either once or at a predefined frequency, receiving consistent service over a planning period that results in time windows and service consistently provided by the same vehicle. The proposed model can reduce the travel times of vehicles in the planning horizon and achieve consistency in arrival times to customers and achieve consistency in the vehicles that visit the customers.

The mathematical model was tested on structured instances to analyse how it would handle some changes to aspects such as customer distribution (uniform or clustered), depot location, and arrival times to customers and whether the removal of certain constraints would affect the performance of the objective function. The model was developed in the AMPL programming language and solved with solvers CPLEX and Gurobi. The results are outstanding for the problem considered in this study.

For future work, we suggest examining the multi-objective problem [35] to minimize both vehicle travel time and the variation in arrival time to customers, extending the model to consider a heterogeneous fleet and real constraints (rich vehicle routing problems), and developing heuristics for medium- and large-scale instances. Future research uncertainty into the model and a stochastic approach such as Sample Average Approximation must be considered to solve the considered problem. Finally, heuristic and metaheuristic algorithms based on works proposed by [36-40] could be implemented to solve large instances of issues.

Acknowledgement

This work has been partially supported by the research project 2060222 IF/R from Universidad del Bio-Bio and the supercomputing infrastructure NLHPC (ECM-02). Also, it is acknowledging the partial support of the Universidad del Valle, Cali, Colombia.

References

- [1] Wu, Q., Wang, X., He, Y.D., Xuan, J., He, W.D. (2018). A robust hybrid heuristic algorithm to solve multi-plant milk-run pickup problem with uncertain demand in automobile parts industry, *Advances in Production Engineering & Management*, Vol. 13, No. 2, 169-178, doi: [10.14743/apem2018.2.282](https://doi.org/10.14743/apem2018.2.282).
- [2] Giaglis, G.M., Minis, I., Tatarakis, A., Zeimpekis, V. (2004). Minimizing logistics risk through real-time vehicle routing and mobile technologies: Research to date and future trends, *International Journal of Physical Distribution & Logistics Management*, Vol. 34, No. 9, 749-764, doi: [10.1108/09600030410567504](https://doi.org/10.1108/09600030410567504).
- [3] Feillet, D., Garaix, T., Lehuédé, F., Péton, O., Quadri, D. (2014). A new consistent vehicle routing problem for the transportation of people with disabilities, *Networks*, Vol. 63, No. 3, 211-224, doi: [10.1002/net.21538](https://doi.org/10.1002/net.21538).
- [4] Somboonwiwat, T., Khompatraporn, C., Miengarrom, T., Lerdluechachai, K. (2018). A bi-objective environmental-economic optimisation of hot-rolled steel coils supply chain: A case study in Thailand, *Advances in Production Engineering & Management*, Vol. 13, No. 1, 93-106, doi: [10.14743/apem2018.1.276](https://doi.org/10.14743/apem2018.1.276).
- [5] Tarantilis, C.D., Stavropoulou, F., Repoussis, P.P. (2012). A template-based tabu search algorithm for the consistent vehicle routing problem, *Expert Systems with Applications*, Vol. 39, No. 4, 4233-4239, doi: [10.1016/j.eswa.2011.09.111](https://doi.org/10.1016/j.eswa.2011.09.111).
- [6] Bernal, J., Escobar, J.W., Paz, J.C., Linfati, R., Gatica, G. (2018). A probabilistic granular tabu search for the distance constrained capacitated vehicle routing problem, *International Journal of Industrial and Systems Engineering*, Vol. 29, No. 4, 453-477, doi: [10.1504/IJISE.2018.094267](https://doi.org/10.1504/IJISE.2018.094267).
- [7] Escobar, J.W. (2014). Heuristic algorithms for the capacitated location-routing problem and the multi-depot vehicle routing problem, *4OR - A Quarterly Journal of Operations Research*, Vol. 12, 99-100, doi: [10.1007/s10288-013-0241-4](https://doi.org/10.1007/s10288-013-0241-4).
- [8] Escobar, J.W., Linfati, R., Toth, P., Baldoquin, M.G. (2014). A hybrid granular tabu search algorithm for the multi-depot vehicle routing problem, *Journal of Heuristics*, Vol. 20, No. 5, 483-509, doi: [10.1007/s10732-014-9247-0](https://doi.org/10.1007/s10732-014-9247-0).
- [9] Schulze, J., Fahle, T. (1999). A parallel algorithm for the vehicle routing problem with time window constraints, *Annals of Operations Research*, Vol. 86, 585-607, doi: [10.1023/A:1018948011707](https://doi.org/10.1023/A:1018948011707).
- [10] Jozefowicz, N., Semet, F., Talbi, E.-G. (2008). Multi-objective vehicle routing problems, *European Journal of Operational Research*, Vol. 189, No. 2, 293-309, doi: [10.1016/j.ejor.2007.05.055](https://doi.org/10.1016/j.ejor.2007.05.055).
- [11] Bolaños, R.I., Escobar, J.W., Echeverri, M.G. (2018). A metaheuristic algorithm for the multi-depot vehicle routing problem with heterogeneous fleet, *International Journal of Industrial Engineering Computations*, Vol. 9, No. 4, 461-478, doi: [10.5267/j.ijiec.2017.11.005](https://doi.org/10.5267/j.ijiec.2017.11.005).
- [12] Gaudioso, M., Paletta, G. (1992). A heuristic for the periodic vehicle routing problem, *Transportation Science*, Vol. 26, No. 2, 86-92, doi: [10.1287/trsc.26.2.86](https://doi.org/10.1287/trsc.26.2.86).
- [13] Min, J.N., Jin, C., Lu, L.J. (2019). Maximum-minimum distance clustering method for split-delivery vehicle-routing problem: Case studies and performance comparisons, *Advances in Production Engineering & Management*, Vol. 14, No. 1, 125-135, doi: [10.14743/apem2019.1.316](https://doi.org/10.14743/apem2019.1.316).
- [14] Sepúlveda, J., Escobar, J.W., Adarme-Jaimes, W. (2014). An algorithm for the routing problem with split deliveries and time windows (SDVRPTW) applied on retail SME distribution activities, *DYNA*, Vol. 81, No. 187, 223-231, doi: [10.15446/dyna.v81n187.46104](https://doi.org/10.15446/dyna.v81n187.46104).
- [15] Bernal, J., Escobar, J.W., Linfati, R. (2017). A granular tabu search algorithm for a real case study of a vehicle routing problem with a heterogeneous fleet and time windows, *Journal of Industrial Engineering and Management*, Vol. 10, No. 4, 646-662, doi: [10.3926/jiem.2159](https://doi.org/10.3926/jiem.2159).
- [16] Stewart Jr., W.R., Golden, B.L. (1983). Stochastic vehicle routing: A comprehensive approach, *European Journal of Operational Research*, Vol. 14, No. 4, 371-385, doi: [10.1016/0377-2217\(83\)90237-0](https://doi.org/10.1016/0377-2217(83)90237-0).
- [17] Dumas, Y., Desrosiers, J., Soumis, F. (1991). The pickup and delivery problem with time windows, *European Journal of Operational Research*, Vol. 54, No. 1, 7-22, doi: [10.1016/0377-2217\(91\)90319-Q](https://doi.org/10.1016/0377-2217(91)90319-Q).
- [18] Santa Chávez, J.J., Escobar, J.W., Echeverri, M.G., Meneses, C.A.P. (2018). A heuristic algorithm based on tabu search for vehicle routing problems with backhauls, *Decision Science Letters*, Vol. 7, No. 2, 171-180, doi: [10.5267/j.dsl.2017.6.001](https://doi.org/10.5267/j.dsl.2017.6.001).
- [19] Cao, Q.K., Yang, K.W., Ren, X.Y. (2017). Vehicle routing optimization with multiple fuzzy time windows based on improved wolf pack algorithm, *Advances in Production Engineering & Management*, Vol. 12, No. 4, 401-411, doi: [10.14743/apem2017.4.267](https://doi.org/10.14743/apem2017.4.267).
- [20] Linfati, R., Escobar, J.W. (2018). Reoptimization heuristic for the capacitated vehicle routing problem, *Journal of Advanced Transportation*, Vol. 2018, Article ID 3743710, doi: [10.1155/2018/3743710](https://doi.org/10.1155/2018/3743710).
- [21] Groër, C., Golden, B., Wasil, E. (2009). The consistent vehicle routing problem, *Manufacturing & Service Operations Management*, Vol. 11, No. 4, 630-643, doi: [10.1287/msom.1080.0243](https://doi.org/10.1287/msom.1080.0243).
- [22] Linfati, R., Escobar, J.W., Escalona, J. (2018). A two-phase heuristic algorithm for the problem of scheduling and vehicle routing for delivery of medication to patients, *Mathematical Problems in Engineering*, Vol. 2018, Article ID 8901873, doi: [10.1155/2018/8901873](https://doi.org/10.1155/2018/8901873).

- [23] Den Ridder, R. (2014). ConRTR, a consistent vehicle routing problem algorithm, from <https://pdfs.semanticscholar.org/3a72/80ff17eaf4938b1fe7f4944e5c352a90d0e6.pdf>, accessed September 12, 2019.
- [24] Lehuédé, F., Pavageau, C., Péton, O. (2008). Un système d'aide à la décision pour planifier les transports vers les établissements médico-sociaux, In: *Proceedings of 5è conférence Handicap*, Paris, France, 168-173.
- [25] Lespay, H., Suchan, K. (2019). The consistent vehicle routing problem for a food distribution firm, In: *Proceedings of Verolog Conference, Workshop of the EURO Working Group on Vehicle Routing and Logistics Optimization*, Seville, Spain, 1-2.
- [26] Campelo, P., Neves-Moreira, F., Amorim, P., Almada-Lobo, B. (2019). Consistent vehicle routing problem with service level agreements: A case study in the pharmaceutical distribution sector, *European Journal of Operational Research*, Vol. 273, No. 1, 131-145, doi: [10.1016/j.ejor.2018.07.030](https://doi.org/10.1016/j.ejor.2018.07.030).
- [27] Macdonald, T., Dörner, K., Gandibleux, X. (2009). Metaheuristics for the consistent nurse scheduling and routing problem, from <http://oro.univ-nantes.fr/sujets-09-10/macdonald.pdf>, accessed September 20, 2019.
- [28] Goeke, D., Roberti, R., Schneider, M. (2019). Exact and heuristic solution of the consistent vehicle-routing problem, *Transportation Science*, Vol. 53, No. 4, 986-1000, doi: [10.1287/trsc.2018.0864](https://doi.org/10.1287/trsc.2018.0864).
- [29] Xu, Z., Cai, Y. (2018). Variable neighborhood search for consistent vehicle routing problem, *Expert Systems with Applications*, Vol. 113, 66-76, doi: [10.1016/j.eswa.2018.07.007](https://doi.org/10.1016/j.eswa.2018.07.007).
- [30] Stavropoulou, F., Repoussis, P.P., Tarantilis, C.D. (2019). The vehicle routing problem with profits and consistency constraints, *European Journal of Operational Research*, Vol. 274, No. 1, 340-356, doi: [10.1016/j.ejor.2018.09.046](https://doi.org/10.1016/j.ejor.2018.09.046).
- [31] Kovacs, A.A., Golden, B.L., Hartl, R.F., Parragh, S.N. (2015). The generalized consistent vehicle routing problem, *Transportation Science*, Vol. 49, No. 4, 796-816, doi: [10.1287/trsc.2014.0529](https://doi.org/10.1287/trsc.2014.0529).
- [32] Kovacs, A.A., Parragh, S.N., Hartl, R.F. (2015). The multi-objective generalized consistent vehicle routing problem, *European Journal of Operational Research*, Vol. 247, No. 2, 441-458, doi: [10.1016/j.ejor.2015.06.030](https://doi.org/10.1016/j.ejor.2015.06.030).
- [33] Miller, C.E., Tucker, A.W., Zemlin, R.A. (1960). Integer programming formulation of traveling salesman problems, *Journal of the ACM*, Vol. 7, No. 4, 326-329, doi: [10.1145/321043.321046](https://doi.org/10.1145/321043.321046).
- [34] Uchoa, E., Pecin, D., Pessoa, A., Poggi, M., Vidal, T., Subramanian, A. (2017). New benchmark instances for the capacitated vehicle routing problem, *European Journal of Operational Research*, Vol. 257, No. 3, 845-858, doi: [10.1016/j.ejor.2016.08.012](https://doi.org/10.1016/j.ejor.2016.08.012).
- [35] Zhao, P.X., Luo, W.H., Han, X. (2019). Time-dependent and bi-objective vehicle routing problem with time windows, *Advances in Production Engineering & Management*, Vol. 14, No. 2, 201-212, doi: [10.14743/apem2019.2.322](https://doi.org/10.14743/apem2019.2.322).
- [36] Arab, R., Ghaderi, S.F., Tavakkoli-Moghaddam, R. (2020). Two efficient meta-heuristic algorithms for the robust inventory routing problem with backhaul, *Tehnički Vjesnik – Technical Gazette*, Vol. 27, No. 3, 793-802, doi: [10.17559/TV-20180814091028](https://doi.org/10.17559/TV-20180814091028).
- [37] Gocken, T., Yaktubay, M. (2019). Comparison of different clustering algorithms via genetic algorithm for VRPTW, *International Journal of Simulation Modelling*, Vol. 18, No. 4, 574-585, doi: [10.2507/IJSIMM18\(4\)485](https://doi.org/10.2507/IJSIMM18(4)485).
- [38] Sepúlveda, J., Escobar, J.W., Adarme-Jaimes, W. (2014). An algorithm for the routing problem with split deliveries and time windows (SDVRPTW) applied on retail SME distribution activities, *DYNA*, Vol. 81, No. 187, 223-231, doi: [10.15446/dyna.v81n187.46104](https://doi.org/10.15446/dyna.v81n187.46104).
- [39] Santa Chávez, J.J., Escobar, J.W., Echeverri, M.G., Meneses, C.A.P. (2018). A heuristic algorithm based on tabu search for vehicle routing problems with backhauls, *Decision Science Letters*, Vol. 7, No. 2, 171-180, doi: [10.5267/j.dsl.2017.6.001](https://doi.org/10.5267/j.dsl.2017.6.001).
- [40] Santa Chávez, J.J., Echeverri, M.G., Escobar, J.W., Meneses, C.A.P. (2015). A metaheuristic ACO to solve the multi-depot vehicle routing problem with backhauls, *International Journal of Industrial Engineering and Management (IJIEEM)*, Vol. 6, No. 2, 49-58.

Decentralized optimization of the flexible production lines

Malega, P.^a, Rudy, V.^a, Kanász, R.^b, Gazda, V.^{c,*}

^aTechnical university of Kosice, Faculty of Mechanical Engineering; Institute of management, Industrial and Digital Engineering, Slovak Republic

^bMarwin Cassovia Soft, Ltd.

^cTechnical university of Kosice, Faculty of Economics, Department of Finance, Slovak Republic

ABSTRACT

The shortening of the production cycle and increasing impact of the technological innovations evokes improvement of the methods used in the production line scheduling. The aim of the presented research is a proposal of the decision model that enables a flexible reaction to the changing production conditions. The central decisions are substituted by the decisions performed on the independent machines level. The machines utilize rather restricted information on the capacities utilization of their technological neighbours. The decisions follow the decision tables (decision chromosomes) re-coded in the course of the evolution. The Genetic algorithms standing behind the model, enable identification of the highly acceptable solutions that is proved by a set of the simulation experiments. The set of independent machines becomes self-organized, having a significant positive effect on the production line capacities utilization. The decentralization aspect makes our proposal somewhat different from other research in the field. The philosophy is built on the fact that each machine makes its production decision based on the available information, which it has at its disposal in a given time. The machine flexible reacts to its neighbours' capacity utilization (machine before and after given production machine) in the production process flow.

© 2020 CPE, University of Maribor. All rights reserved.

ARTICLE INFO

Keywords:

Production line;
Job shop problem (JSP);
Decentralised optimization;
Production scheduling;
Shortest processing time rule;
Self-organization;
Genetic algorithm;
Decision table

*Corresponding author:

vladimir.gazda@tuke.sk
(Gazda, V.)

Article history:

Received 15 May 2020
Revised 16 October 2020
Accepted 19 October 2020

1. Introduction

Since the first mathematical formulations of production scheduling in the late 1950s, the topic became a subject of intensive scientific discourse. The problem formulations involve the assignment of scarce resources, typically machines, to the competing tasks over time to optimize some characteristics of system performance. Proper production scheduling promises considerable economizing of the production line performance achieved without paying any additional costs.

Development of operational research methods and growth of computational capacities significantly modified the character of the problem settings, covering the formulation of the dynamic programming, integer programming, and various types of the heuristic algorithms until the end of the 90s (for more detailed preview see [1]). However, the rapid development of the technologies, increasing production complexity and shortening of the production cycles limit the potential of central control of the production lines. Designing flexible production lines capable of reflecting sudden unexpected changes in the environment becomes inevitable. In this framework, the theory of Artificial Intelligence [2] offers the methods feasible for utilization in the new approach to production scheduling. Notably, the methodology of Agent based modelling becomes quite popular. Agents constitute the independent decision units enjoying simple behaviour rules

and following their objectives. Their mutual interactions lead to the synchronization of the whole system, often generating entirely unexpected macro-features – called emerging events. Workplaces (or machines)¹, playing the role of the agents, could be successfully utilized in the proposals of the new flexible production line scheduling. We assume that the self-organization of the workplaces enables a flexible reaction to the unexpected production shocks and quite stable solutions with regard to the changing environment.

The paper presents the application of the job shop scheduling based on the autonomous working place decisions utilizing the evolutionary mechanism of genetic algorithms. We show that even simple decision rules generated in the process of the genetic evolution lead to systematic shortening of the jobs makespan and thus economizing of the production lines. The simulation results are compared with the traditional approach to scheduling – in our case, the shortest processing time (SPT) rule. We show that the proposed approach provides computationally stable solutions.

The paper is organized as follows. The second section describes the literature motivating the presented research. The proposed model is introduced in the third section. The fourth section presents the simulation results, and the fifth section discusses the contributions of the paper.

2. Literature review

Different types of scheduling problems vary according to their organization, range, and other details. One of the most discussed types is the Job-shop problem (JSP), where a set of machines processes a group of jobs. Every job consists of a sequence of consecutive operations (tasks), while particular machines perform them in a reserved way. The scheduling constitutes in the assignment of the operations on the machines by optimizing a specific indicator [3]. The most traditional optimization criterion is a makespan, which is the time needed for processing all jobs in a given production schedule.

Besides the minimization of the makespan, the classical JSP also targets improving production efficiency, reduction of the production costs [4], or reduction of the energy consumption [5]. At the same time, it should take into consideration all the constraints known in advance [6]. When speaking about methods, mathematical models and optimizing algorithms dominated in the research published in the last decades of the 20th century [1].

The job-shop scheduling problem (JSP) belongs by its nature to the combinatorial optimization problems [7]. Garey *et al.* and Blazewicz *et al.* pointed to its computational complexity that is NP-hard [8, 9]. Besides, most production systems operate in a dynamic environment where a few unpredictable real-time events can entirely disrupt the previously defined planned schedule. Examples of such real-time events include machine failures, the arrival of imperative jobs, due date changes, and others [10].

Mohapatra *et al.* [11] used a multi-objective approach to solve the planning and scheduling problem. Three different objectives considered in this work are the following: minimisation of (i) makespan, (ii) machining costs, and (iii) idle time of machines. To solve this integration problem, the authors propose an improved controlled elitist non-dominated sorting genetic algorithm to take into account the computational intractability of the problem.

Supsomboon and Vajasuvisom [12] proposed the alternative solutions to improve machinery production process in job shop scheduling using alternative strategies. The authors apply basic principles of operations management. Simulation models showed the great performance to help in making a decision in their real system.

Ma *et al.* [13] developed a novel scheduling method based on the fuzzy satisfaction rate and differential evolution algorithm. In the method, at first the fuzzy parameters are deduced based on normal distribution to calculate the satisfaction rate. After that a differential evolution is proposed.

¹ Terms workplace and machine are used as synonyms in the text

Ojstersek *et al.* [14] emphasises the advantage of a modular adaptive design allowing adaptation of the simulation model to a wide range of multi-objective optimisation problems. The presented MOHKA method is suitable for optimisation results' interactivity between the mathematical and simulation models. The interactive method pointed on the advantage of transferring the optimisation results via a simulation model to a real environment.

Shen and Yao's [15] contribution is the construction of a dynamic multi-objective optimization model for multi-objective dynamic flexible job shop scheduling. Another contribution of this research is the design of a dynamic decision making procedure.

On the other hand, neither the mathematical programming in its classical form nor the intuitive heuristics are capable of tackling the high complexity problems. On the other hand, at the end of the 20th and start of the 21st century, many new computational approaches emerged.

Meolic and Brezocnik [16] propose a novel method for generating and counting solutions of a flexible job shop scheduling problem. They represent the feasible solutions as cubes of a combination set. The algorithm is implemented as a sequence of operations in unique cube *set algebra*.

Yu *et al.* [17] propose a novel two-phase integration of process planning and scheduling approach. The preplanning phase generates a process network for each job with consideration of the static shop floor status. After achieving this goal, the final planning phase simultaneously creates the process plan of each job and the scheduling plan according to the current shop floor status.

The one, dominating the presented research, is the genetic algorithm (GA) method first formulated by J. Holland in his seminal works [18, 19]. However, GA applications in JSP scheduling has a long history starting with Bierswirth *et al.* [20] who published seemingly the first, however, mostly neglected, work.

Janes *et al.* [21] present an efficient genetic algorithm for solving job-shop scheduling problems. In comparison with traditional genetic algorithm, selection and crossover operators were modified. Authors demonstrate the viability of the proposed method for real-world problems and suggest its readiness for application in industry.

Chen and Hao [22] apply the non-dominated sorting genetic algorithm in combination with the multi-objective optimization. An excellent review of the scheduling under Industry 4.0 is provided by Zhang *et al.* [23].

In production scheduling, genetic algorithms (GA) represent schedules as individuals or a population's members. Every individual has its own fitness value. It is measured based on the objective function. This methodical approach works iteratively, and this iteration is a generation. Every generation has individuals and they are from the previous generations. Typically, the population size stays consistent starting with one generation to the next generation.

3. Proposed model

Job Shop Problem (JSP) is very famous optimization problem in software engineering and operational research. The focus is on assigning jobs to resources at particular times. A Job Shop is a work area where various universally useful workstations exist and are utilized to carry out a variety of jobs.

The proposed model assumes that the complexity of the production line prevents its effective centralized scheduling. Instead, the autonomous decision of the individual workplaces (machines) equipped with artificial intelligence is assumed. Similar to some bioinspired algorithms (flocking is the best example), the individual machines react to their neighbours to coordinate their job selection decisions. Thus the production line becomes self-organized and decentralised demonstrating a rather high level of synchronization.

3.1 Basic concepts and definitions

We assume a production line given as a set $M = \{1, 2, \dots, m\}$ of the dedicated machines ordered according to technology process. If $i < k$ ($i, k \in M$) then machine i precedes machine j in the technological flow. There are no machines running in parallel. Set $J = \{1, 2, \dots, n\}$ denotes the

scheduled jobs (i.e., products). Function $t : J \times M \rightarrow N_+$ assigns task processing time $t_{i,j}$ of the i -th job to the machine j .

Operational imbalances of the time schedules imply the accumulation of the semi-products in each machine buffer store. It is fulfilled with the semi-products waiting for the follow-up of the production process. The machine decision consists of the selection and processing of one particular product among the stored semi-products. The machine interrupts production if it has no products in its buffer store. Any interruption prolongs the makespan of the whole production line.

Let us explain the situation of the i -th machine in detail. The i -th machine is preceded by the machine $i - 1$ and succeeded by the machine $i + 1$ in the technological flow. Because of the expression simplicity, the machine indexes are further substituted by $-1, 0, +1$. Machines do not dispose of the complete production line information and know just the content of the buffer store of both its own and the ones of -1 and $+1$ neighbours. The agent does not take into consideration the information regarding tasks, which have already been performed.

The useful information of the buffer store contents (see Eq. 1) is formally expressed as an array of the task capacity demands as follows

$$A^{(k)} = [a_{i,j}^{(k)}] \quad i \in J; \quad j, k \in \{-1, 0, 1\}; \quad j \geq k \tag{1}$$

where (k) denotes the buffer store of the k -th machine (predecessor (-1) , current machine (0) , and its successor (1)); i denotes i -th job; ordered pair (i, j) denotes the task, i.e. the i -th job time capacity requirement against the j -th machine.

An example of the information disposable for one machine is given in Table 1.

Each machine makes its independent decision on which product from its buffer store to choose for further processing. In such a decision situation, the traditional views on the production scheduling would offer mostly various variants of the Shortest Processing Time (SPT) heuristic. It assumes that if the machine is free, then it chooses the task with the smallest processing time requirements on the current machine. However, there are many objections to the effectiveness of this heuristic [24].

Table 1 Example of the buffer stores – arrays $A^{(-1)}, A^{(0)}, A^{(1)}$

Job (i)	$A^{(-1)}$			Job (i)	$A^{(0)}$		Job (i)	$A^{(1)}$
	$A_{i,-1}^{(-1)}$	$A_{i,0}^{(-1)}$	$A_{i,1}^{(-1)}$		$A_{i,0}^{(0)}$	$A_{i,1}^{(0)}$		$A_{i,1}^{(1)}$
1	7	6	8	5	1	2	8	5
3	5	3	2	6	3	1	9	2
2	1	7	1	4	8	4		
				7	6	6		

3.2 Genetic algorithm in strategy decision

The current state of artificial intelligence development offers a wide variety of inspirations for the SPT rule extensions. Highly promising seems to be a J. Koza's ([25]) theory of the genetic programming that is capable of deriving productive decision trees to solve rather complicated and in-formalized decision situations. It also remains the main inspiration for the proposed model. On the other hand, the substitution of the decision trees by rather simple decision tables significantly clarifies the decision logic and reduces the computational capacities requirements. Besides, we consider the deep decision trees to be less stable in the decision rules evolution.

Introduce the row vector-chromosome of the production line (see Eq. 2)

$$x = [x^1 | x^2 | \dots | x^m] \tag{2}$$

that constitutes from the individual machine row vector-chromosomes. The chromosome of the i th machine (see Eq. 3) is defined as

$$x^i = [x_1^i, x_2^i, \dots, x_p^i, x_{p+1}^i, \dots, x_r^i,] \tag{3}$$

where its first p elements encode the simple logical conditions derived from the disposable information (i.e. $A^{(-1)}, A^{(0)}, A^{(1)}$ arrays). The last $r - p$ vector elements encode the product conditional selection strategy based on fulfilment given logical conditions.

All the vector chromosome elements (*genes*) are represented by the trits, which are 3-state analogues to the 2-state bits. Each gene contains one of $-1, 0, 1$ value (*allele* in genetics). The decision table with varying r bits enables modelling 3^r logical condition variants (see Table 2).

First $r - p$ trits have the following interpretation:

- if the gene gets number 1, then the corresponding logical condition is fulfilled;
- if the gene gets value -1 , then the negation of the corresponding logical condition is fulfilled;
- if the gene gets value 0, the validity of the logical condition is inconclusive and is not taken into consideration.

Table 2 The one-machine chromosome vector composition – Example

Trit (gene)	Logical condition	Indicator (allele)
1	$\sum_i a_{i,1}^{(1)} > \max_i a_{i,0}^{(0)}$	
2	$\sum_i a_{i,0}^{(0)} > \max_i a_{i,-1}^{(-1)}$	
3	$\min_i (a_{i,1}^{(1)}) > \min_i (a_{i,0}^{(0)})$	
4	$\text{med}_i (a_{i,1}^{(1)}) > \max_i (a_{i,1}^{(0)})$	
5	$\min_i (a_{i,-1}^{(-1)}) < \min_i (a_{i,0}^{(0)})$	{-1,0,1}
6	$\max_i (a_{i,-1}^{(-1)}) > \max_i (a_{i,0}^{(0)})$	
7	$\sum_i a_{i,-1}^{(-1)} > \sum_i a_{i,0}^{(0)}$	
8	$\max_i (a_{i,0}^{(0)}) > \max_i (a_{i,1}^{(1)})$	
9	$\text{med}_i (a_{i,-1}^{(-1)}) > \max_i (a_{i,0}^{(0)})$	
Trit	Conditional Strategy	Indicator (allele)
10	<pre> if (cond.(1.-9.) == Ind.(1.-9.)) then { if (Ind(10.) == 1) then {choose arg max_i (a_{i,0}^{(0)})} else {if (Ind(10.) == -1) then {choose arg min_i (a_{i,0}^{(0)})} else {evaluate condition Ind(11.)} } } </pre>	{-1,0,1}
11	<pre> if (Ind(11.) == 1) then {choose arg max_i (a_{i,0}^{(0)} + (a_{i,1}^{(1)}))} else {if (Indicator == -1) then {choose arg med_i (a_{i,0}^{(0)})} else {choose arg min_i (a_{i,0}^{(0)} + (a_{i,1}^{(1)}))} } </pre>	{-1,0,1}

Notes: Trit means three-state analogy to the double-state bit. The genes trits contents (alleles) are set in the course of the genetic evolution. First nine trits code confront the logical condition: fulfilled (1); fulfilled negation (-1); not in consideration (0). Trits 10 and 11 code the conditional strategy of the product selection.

In case of the remaining $r - p$ trits, the conditional selection of a strategy is based both on the $-1, 0, 1$ corresponding trit content, and logical fulfilment of the selected conditions from the first r ones.

We assume the strategy space should contain just a few strategies to keep the machine strategy simple and transparent. In an example, we assume the selection of the i^* -th product for further processing based on the following:

- $i^* = \arg \max_i (a_{i,0}^{(0)})$
- $i^* = \arg \min_i (a_{i,0}^{(0)})$
- $i^* = \arg \max_i (a_{i,0}^{(0)} + a_{i,1}^{(1)})$
- $i^* = \arg \min_i (a_{i,0}^{(0)} + a_{i,1}^{(1)})$
- $i^* = \arg \text{med}_i (a_{i,0}^{(0)} + a_{i,1}^{(1)})$

An example of the decision table with $p = 9$ and $r - p = 2$ elements is given in Table 2. In the following, the simulation results section provides an analysis and simulation results of the given example.

4. Simulation results and discussion

Genetic algorithms simulate the Darwinian and Mendelian principles of the evolution that are based on the natural selection of the most fitted individuals and their ability to accommodate to the changing competitive environment. The most fitted individuals (elite) are selected for breeding transferring their genetic information encoded in chromosomes onto the next generation. The rest of the population is assumed to die off and being replaced by the offspring of the breeding parents. The inheritance process is based on the crossing and mutation. While crossing directly transfers the chromosome crossing combination (inheritance information) from the parents onto the next population, the mutation introduces to the population a new piece of information, regardless of whether positive or negative. It has been proved that large and most heterogeneous populations provide best fitted individuals both in case of nature as well as in the case of the artificial simulations.

Table 3 Parameters settings and their interpretation

Parameter	Values	Interpretation
No of machines	10	Dedicated machines ordered according to the technological flow. No machine works in parallel.
No of jobs	30	
TPT	Unif (1,20)	Task processing times (TPTs) are uniformly distributed on the $\{1,2,\dots,20\}$ set.
Prob. of crossing	0.8	When a pair of the machines is selected from the elite category, probability of their breeding is 0.8.
Population	100, 300, 500	Population consists of the set of the virtual production lines each endowed with its specific chromosome. Gene values $(-1, 0, 1;$ alleles) are set randomly according to the uniform distribution in the start of the simulation.
Elitism	0.1,0.2,...0.9	The share of the most fitted individuals selected for breeding and surviving into the next simulation round.
Type of crossing	1-point	The parents' chromosomes cross in one randomly selected position (gene).
Mutation rate	0.01,0.05,0.1	Mutation is based on random flipping of the contents of one randomly selected gene (allele).
End of simulations	50 periods	One period consists of consecutive steps: (1) determination of the elite; (2) crossing; (3) mutation of the offspring; (4) computation of fitness.
Fitness	Negative value of makespan	Fitness is given as a negative value of the makespan.

Simulation models are sensitive to the parameters setting. If set incorrectly, the simulation results usually fail. In case of the proposed model, a few pilot simulations were performed to identify an admissible parameters space worthy for the consideration as given in Table 3.

4.1 Optimal population size, mutation rate and elitism

Determination of the satisfactory population size seems to be crucial for identification of some kind of sub-optimal solution in an acceptable computing time. If one machine-chromosome contains 11 three-state genes (trits), there are $3^{11} = 177147$ possibilities how to code it. Complexity of the problem significantly arises, if considering the machine-line chromosome (if 10 machines, then 3^{110} permutations). In case the population is rather large and heterogeneous, it can contain enough information to achieve the optima by simple parents' chromosomes crossing. On the other hand, a small population may be rather depleted of the genetic content variability and the only chance to converge to some acceptable solutions is introducing of the new information via mutation, but lasting longer. A similar dilemma faces the proper determination of the elitism size. The small one bears the threat of losing the genetic variability causing convergence into a sub-optimal state located far from the optimal one. Alternatively, too large elitism transfers rather ineffective decision strategies to the future generations that can slow the convergence to the needed solutions. If one is speaking about mutation rates, the following is valid in general. As all the mutations are highly random, their effect on searched optima can become rather unambiguous. However, mutation enriches the elite population with the new information, but can also add too much noise and so to prevent effective stabilization of the convergent results. Fig. 1 and Fig. 2 depict the situation that outlines the fixation of the parameters for the following analysis as follows: mutation rate 0.05, population 300, elitism 0.7.

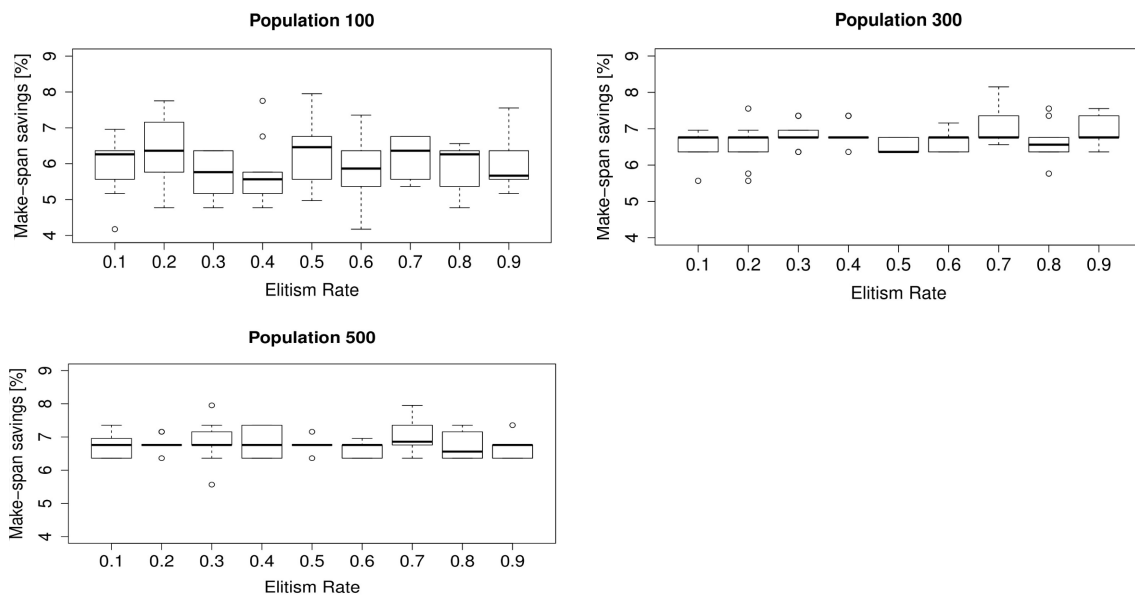


Fig. 1 Makespans distribution by different populations and elitism rates – 10 simulations (mutation rate 0.05; time savings against the Shortest processing time strategy)

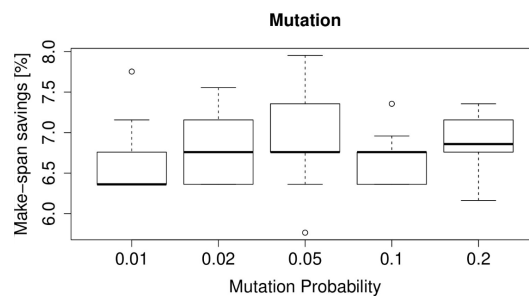


Fig. 2 Makespans distribution – 10 repeating simulations run by different mutation rates (elitism rate = 0.7; time savings against the Shortest processing time strategy)

4.2 Sensitivity to initial conditions, dynamics, and evolutionary stability

Any application of the GA deals with the stochasticity regarding (i) the initial population random settings; and (ii) random character of the mutations and crossing. It may lead to inconsistent computational results in repeating simulations, which makes performing of the computational stability analysis quite inevitable. In addition, the following research includes a simple investigation of the evolutionary dynamics to identify typical patterns in the evolutionary trajectory.

Following the principle of the previous analysis based on the repeating simulations using the same set of jobs and machines, 10 repeating simulation runs provide the results depicted in Fig. 3.

The genetic algorithm improves the autonomous decision mechanism of the self-organized machines in the first two simulation rounds rather significantly. One can distinguish two kinds of the simulation evolution. Smooth evolution process starts in the third iteration round lasting until the sixth one. This phase is typical with a quite small improvement of the makespan shortening. In the seventh round, the phase transition happened in the majority of the simulations. The makespan savings suddenly jumped by about 2 percent on the average. Then, the smooth evolution starts again and follows until the ninth round. After that, the phase transition happened again and the renewed machine decision rules demonstrate relevant improvements as the makespans significantly shortened on the average. The simulation round 15 starts the smooth evolutionary period, again, lasting with insignificant improvements until the end of the simulation, i.e. the 50th simulation round. The system reached its evolutionary equilibrium.

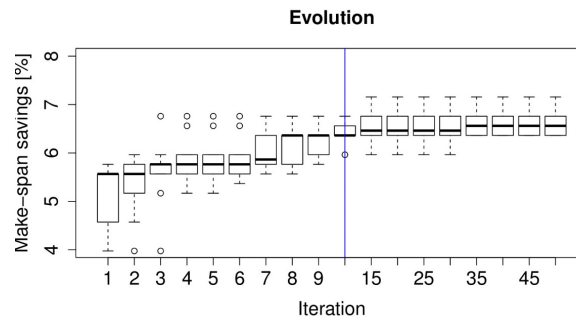


Fig. 3 Time evolution of the makespans savings – Boxplots of ten repeating simulations (time savings against the Shortest processing time strategy)

4.3 Changing list of the jobs – Robustness analysis

The above introduced analysis depicts the results of repeating simulations using the same list of jobs. On the other hand, one can object to the approach universality and its robustness against changing jobs list composition. To reduce the doubts, five random job list composition was generated (10 machines, 30 jobs, changing task requirements). The simulation parameters remain the same, i.e., elitism rate 0.8; mutation rate 0.05; no of iterations 50, population size 300.

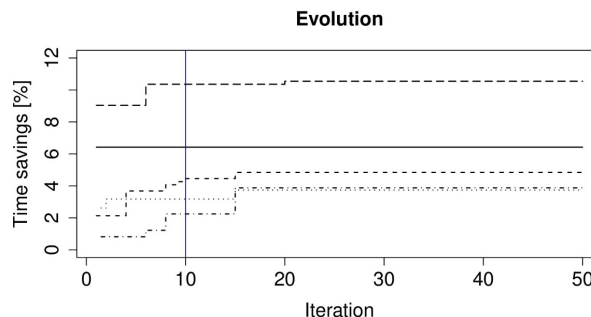


Fig. 4 The best makespans savings in case of 5 independent simulations with changing job list compositions (time savings against the Shortest processing time strategy)

Fig. 4 introduces the evolution of the best makespans within each of the five simulations. If compared the evolution led to the typical patterns given in previous analysis (see previous subsection and Fig. 3), evolution in the framework of the previously identified evolutionary pattern is obvious. The most significant changes happened in the first 10 simulation rounds alternating periods of the phase transition and quite stable development. The significant improvements of the best makespans were concentrated into the first 15 simulation rounds. On the other hand, there is the case of one simulation run, which did not demonstrate any improvement of the best decision during the whole simulation. It is possible if the randomly composed initial population contains at least one rather good production line chromosome that dominates during the whole evolutionary process. On the other hand, we consider this situation as considerably rare.

5. Conclusion

The technological complexity, shortening of the production cycles, and increasing specialization place high claims on the production planning and control of the production lines. At the same time, increasing of the IKT capacities and introduction of the artificial intelligence methods create new space for the yet unexperienced opportunities.

The presented research can be interpreted, of course, in the sense of a direction into a machine-to-machine economy (M2M economy). M2M economy means that machines are part of a system of interconnected machines without requiring human-to-human or human-to-machine interaction. At the same time, the machines are considered as autonomous participants. The simulation results prove that autonomous machines grouped in the production line can react flexibly to conditions arising in a given situation. In our case, every machine evaluates its situation according to its neighbours (predecessors and successors) and is independent in this decision-making process.

The presented research proposes a new form of the production line scheduling utilizing the Genetic algorithm methods in the framework of the machine line constituted from the independently deciding machines. Each machine makes its production decision based on the available somewhat restricted information; but flexible reacting to the capacity utilization of its neighbours in the technological process.

The presented research includes the details of the model proposal, model calibration, and analysis of the parameter settings. Sensitivity to the initial conditions is also included.

Despite using a quite simplifying example case, the proposed approach demonstrates a significant shortening of the makespans if compared to the Shortest processing time benchmark.

Acknowledgement

The paper was supported by the KEGA 002TUKE-4/2020, APVV-15-0358 and APVV-18-0368 projects.

References

- [1] Maccarthy, B.L., Liu, J. (1993). Addressing the gap in scheduling research: A review of optimization and heuristic methods in production scheduling, *International Journal of Production Research*, Vol. 31, No. 1, 59-79, [doi: 10.1080/00207549308956713](https://doi.org/10.1080/00207549308956713).
- [2] Çaliş, B., Bulkan, S. (2015). A research survey: Review of AI solution strategies of job shop scheduling problem, *Journal of Intelligent Manufacturing*, Vol. 26, No. 5, 961-973, [doi: 10.1007/s10845-013-0837-8](https://doi.org/10.1007/s10845-013-0837-8).
- [3] Nouiri, M., Bekrar, A., Jemai, A., Niar, S., Ammari, A.C. (2018). An effective and distributed particle swarm optimization algorithm for flexible job-shop scheduling problem, *Journal of Intelligent Manufacturing*, Vol. 29, No. 3, 603-615, [doi: 10.1007/s10845-015-1039-3](https://doi.org/10.1007/s10845-015-1039-3).
- [4] Hoogeveen, H. (2005). Multicriteria scheduling, *European Journal of Operational Research*, Vol. 167, No. 3, 592-623, [doi: 10.1016/j.ejor.2004.07.011](https://doi.org/10.1016/j.ejor.2004.07.011).
- [5] Liu, J., Luo, X.-G., Zhang, X.-M., Zhang, F., Li, B.-N. (2013). Job scheduling model for cloud computing based on multi-objective genetic algorithm, *International Journal of Computer Science Issues*, Vol. 10, No. 1, 134-139.
- [6] Zhang, G., Shao, X., Li, P., Gao, L. (2009). An effective hybrid particle swarm optimization algorithm for multi-objective flexible job-shop scheduling problem, *Computers & Industrial Engineering*, Vol. 56, No. 4, 1309-1318, [doi: 10.1016/j.cie.2008.07.021](https://doi.org/10.1016/j.cie.2008.07.021).

- [7] Sonmez, A.I., Baykasoglu, A. (1998). A new dynamic programming formulation of $(n \times m)$ flowshop sequencing problems with due dates, *International Journal of Production Research*, Vol. 36, No. 8, 2269-2283, doi: [10.1080/002075498192896](https://doi.org/10.1080/002075498192896).
- [8] Garey, M.R., Johnson, D.S., Sethi, R. (1976). The complexity of flowshop and jobshop scheduling, *Mathematics of Operations Research*, Vol. 1, No. 2, 117-129, doi: [10.1287/moor.1.2.117](https://doi.org/10.1287/moor.1.2.117).
- [9] Błażewicz, J., Domschke, W., Pesch, E. (1996). The job shop scheduling problem: Conventional and new solution techniques, *European Journal of Operational Research*, Vol. 93, No. 1, 1-33, doi: [10.1016/0377-2217\(95\)00362-2](https://doi.org/10.1016/0377-2217(95)00362-2).
- [10] Ouelhadj, D., Petrovic, S. (2009). A survey of dynamic scheduling in manufacturing systems, *Journal of Scheduling*, Vol. 12, No. 4, 417-431, doi: [10.1007/s10951-008-0090-8](https://doi.org/10.1007/s10951-008-0090-8).
- [11] Mohapatra, P., Nayak, A., Kumar, S.K., Tiwari, M.K. (2015). Multi-objective process planning and scheduling using controlled elitist non-dominated sorting genetic algorithm, *International Journal of Production Research*, Vol. 53, No. 6, 1712-1735, doi: [10.1080/00207543.2014.957872](https://doi.org/10.1080/00207543.2014.957872).
- [12] Supsomboon, S., Vajasuviwon, A. (2016). Simulation model for job shop production process improvement in machine parts manufacturing, *International Journal of Simulation Modelling*, Vol. 15, No. 4, 611-622, doi: [10.2507/IJSIMM15\(4\)3.352](https://doi.org/10.2507/IJSIMM15(4)3.352).
- [13] Ma, D.Y., He, C.H., Wang, S.Q., Han, X.M., Shi, X.H. (2018). Solving fuzzy flexible job shop scheduling problem based on fuzzy satisfaction rate and differential evolution, *Advances in Production Engineering & Management*, Vol. 13, No. 1, 44-56, doi: [10.14743/apem2018.1.272](https://doi.org/10.14743/apem2018.1.272).
- [14] Ojstersek, R., Lalic, D., Buchmeister, B. (2019). A new method for mathematical and simulation modelling interactivity: A case study in flexible job shop scheduling, *Advances in Production Engineering & Management*, Vol. 14, No. 4, 435-448, doi: [10.14743/apem2019.4.339](https://doi.org/10.14743/apem2019.4.339).
- [15] Shen, X.-N., Yao, X. (2015). Mathematical modeling and multi-objective evolutionary algorithms applied to dynamic flexible job shop scheduling problems, *Information Sciences*, Vol. 298, No. 20, 198-224, doi: [10.1016/j.ins.2014.11.036](https://doi.org/10.1016/j.ins.2014.11.036).
- [16] Meolic, R., Brezočnik, Z. (2018). Flexible job shop scheduling using zero-suppressed binary decision diagrams, *Advances in Production Engineering & Management*, Vol. 13, No. 4, 373-388, doi: [10.14743/apem2018.4.297](https://doi.org/10.14743/apem2018.4.297).
- [17] Yu, M.R., Yang, B., Chen, Y. (2018). Dynamic integration of process planning and scheduling using a discrete particle swarm optimization algorithm, *Advances in Production Engineering & Management*, Vol. 13, No. 3, 279-296, doi: [10.14743/apem2018.3.290](https://doi.org/10.14743/apem2018.3.290).
- [18] Holland, J.H. (1973). Genetic algorithms and the optimal allocation of trials, *SIAM Journal on Computing*, Vol. 2, No. 2, 88-105, doi: [10.1137/0202009](https://doi.org/10.1137/0202009).
- [19] Holland, J.H., Reitman, J.S. (1978). Cognitive systems based on adaptive algorithms, *Pattern-Directed Inference Systems*, Vol. 1, No. 1, 313-329, doi: [10.1016/B978-0-12-737550-2.50020-8](https://doi.org/10.1016/B978-0-12-737550-2.50020-8).
- [20] Bierwirth, C., Kopfer, H., Mattfeld, D.C., Rixen, I. (1995). Genetic algorithm based scheduling in a dynamic manufacturing environment, In: *Proceedings of the Second Conference on Evolutionary Computation*, Perth, Australia, 439-443, doi: [10.1109/ICEC.1995.489188](https://doi.org/10.1109/ICEC.1995.489188).
- [21] Janes, G., Perinic, M., Jurkovic, Z. (2017). An efficient genetic algorithm for job shop scheduling problems, *Tehnički Vjesnik – Technical Gazette*, Vol. 24, No. 4, 1243-1247, doi: [10.17559/TV-20150527133957](https://doi.org/10.17559/TV-20150527133957).
- [22] 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\(4\)C017](https://doi.org/10.2507/IJSIMM17(4)C017).
- [23] Zhang, J., Ding, G., Zou, Y., Qin, S., Fu, J. (2019). Review of job shop scheduling research and its new perspectives under Industry 4.0, *Journal of Intelligent Manufacturing*, Vol. 30, 1809-1830, doi: [10.1007/s10845-017-1350-2](https://doi.org/10.1007/s10845-017-1350-2).
- [24] Oral, M., Malouin, J.-L. (1973). Evaluation of the shortest processing time scheduling rule with truncation process, *AIIE Transactions*, Vol. 5, No. 4, 357-365, doi: [10.1080/05695557308974923](https://doi.org/10.1080/05695557308974923).
- [25] Koza, J.R. (1992). *Genetic programming: On the programming of computers by means of natural selection*, MIT press, Cambridge, USA.

Interactive impacts of overconfidence and fairness concern on supply chain performance

Zhang, Z.J.^a, Wang, P.^{b,*}, Wan, M.Y.^c, Guo, J.H.^a, Luo, C.L.^c

^aSchool of Transportation and Logistics, East China Jiaotong University, Nanchang, P.R. China

^bDongwu Business School, Soochow University, Suzhou, P.R. China

^cSchool of Information Management, Jiangxi University of Finance and Economics, Nanchang, P.R. China

ABSTRACT

For exploring the interactive impacts of overconfidence and fairness concern on optimal decisions of manufacturer and retailer, we establish Stackelberg models with these two behavioural preferences in a two-echelon supply chain, wherein retailer has two behavioural preferences. The optimal equilibrium results are compared in different scenarios, namely the retailer with no behavioural preference, with single-behavioural preference and with the two behavioural preferences. Although previous literatures have proven that overconfidence or fairness concern has a negative influence on retailer, we find that the retailer always benefits from these two behavioural preferences, whether it is retail price, sales effort or utility. This is because when the overconfident degree is within a reasonable range, overconfidence and fairness concern have a positive influence on retailer's decision-making, and when the overconfident degree is high, the fairness concern preference can suppress the adverse effects caused by overconfidence. Compared with the preference of fairness concern, the overconfident preference plays a leading role in supply chain performance, which mainly manifests in retailer's decisions, utility and manufacturer profit. Moreover, the wholesale price and profit of the rational manufacturer may become worsen due to the fairness concern of retailer.

© 2020 CPE, University of Maribor. All rights reserved.

ARTICLE INFO

Keywords:
Supply chain;
Supply chain management;
Modelling;
Performance;
Overconfidence;
Fairness concern;
Behavioural operation;
Stackelberg game

**Corresponding author:*
wangpeng_wl@126.com
(Wang, P.)

Article history:
Received 10 September 2020
Revised 8 October 2020
Accepted 12 October 2020

1. Introduction

Many firms and individuals often show characteristics of bounded rationality in decision-making process, such as overconfidence and fairness concern [1]. The two behavioural preferences are observed anecdotally in practice and coexist in the decision process. Before the onset of Double 11 of 2014, a CEO of Taobao predicts that the return rate of commodity may be in single-digit percentage points, yet the real return rate is 69 %, and the complaint rate is higher than usual, such as Haier complaint rate is 54.2 %¹. The huge discount induces customers to make irrational shopping, leading to an augment in return orders and complaint rates. Moreover, the prediction of CEO and the high return rate of consumer reveal that many decision makers are overconfident. Consequently, wrong decisions cause profit losses and merchants concern fairness issues about revenue allocation, resulting in a high complaint rate [1]. Moreover, overconfidence and fairness concern simultaneously affect retailer's decisions in supply chain. For instance, in "Dou-

¹<https://www.chinainternetwatch.com/11643/false-prosperity-behind-double-11-refund-rate-increased-to-69/>.

ble 11" shopping festival, a retailer believes that he has better marketing capabilities, and thus orders more products from a supplier. Nevertheless, he obtains poor profit due to excessive inventories and marketing expenses. Meanwhile, because of an unfairly disproportionate share of the profit between the retailer and the supplier, he reduces order quantities and increases retail price [2].

Fairness concern is an important factor when the revenue allocation is unfair in a supply chain system [3], and overconfidence is a commonly type of irrational behaviour when retailer overestimates the influence of his sales effort level on market demand. Although the effects of overconfidence and fairness concern on decisions and profits have been extensively studied, the existing studies just consider the overconfidence or the fairness concern in supply chain. But study simultaneously considering the two behaviour preferences for one decision maker is lacking. Many researchers think overconfidence and fairness concern result in self-harm, and are not good for the whole supply chain system [4, 5]. Moreover, social preference, i.e., fairness concern, can be regarded as an important complement to self-regarding preference, i.e., overconfidence [6]. In practice, when a retailer has both the two behavioural preferences, he may benefit from these preferences. Therefore, we attempt to investigate the effects of overconfidence and fairness concern on decisions, profits and utilities in a two-echelon supply chain.

The following study questions should be answered, First, how do the retailer's fairness concern and overconfidence affect the pricing and sales effort for the two-echelon supply chain? Second, what are the impacts of the two behavioural preferences on the profitability and utility of supply chain participants? Third, in the presence of fairness concern and overconfidence, what is the main behavioural preference affecting decision variables and profit?

To investigate these questions, we consider a two-echelon supply chain consisting of a rational manufacturer and a retailer. The assumption that the manufacturer is the leader and the retailer is the follower is proposed, because many manufacturers in the automotive industry or other industries (e.g., Ford, Toyota, Dell and Lenovo) have the power to set the wholesale price, whereas retailers must follow the manufacturer's strategy and accept the price, and then determine a selling price and sales effort level. However, a retailer usually overestimates the effect of sales effort on demand, and pays attention to whether the distribution of profit is fair between participants in a supply chain. Thus, we consider three critical scenarios: (a) retailer with no behavioural preferences, wherein the retailer is rational; (b) retailer with single-behavioural preference, wherein the retailer has a preference, i.e., overconfidence or fairness concern; (c) retailer with two behavioural preferences, wherein the retailer has both overconfidence and fairness concern. Then, we construct three Stackelberg game models in different scenarios, respectively. Specially, we analyse the effect of two behavioural preferences on prices, sales effort, profitability and utility of supply chain members.

The rest of the paper is organized as follows. In section 2, we introduce the literature review. Following that, we describe the basic Stackelberg game model under the retailer with unbounded rationality, and extend the model in which the retailer has the two behavioural preferences; the two models are then solved. In Section 4, we discuss the effect of behavioural preferences in the supply chain under three scenarios. Section 5 presents numerical analysis to examine the theoretical models and propositions. Section 6 concludes and provides direction for future research.

2. Literature review

Two streams of literature research are relevant to the study: overconfidence preference and fairness concern preference in supply chain.

2.1 Overconfidence preference in supply chain

Overconfidence is a common cognitive bias in which decision maker often overestimates his judging ability [7]. This preference is widely existed in behavioural operations management. Ren *et al.* [8] introduce overconfidence in supply chain to study the decisions of overconfident newsvendor, then demonstrate that overconfidence significantly correlates with order bias by

introducing the debiasing technique in an experimental study. Kirshner *et al.* [9] think that greater overconfidence typically results in lower margins, higher inventory and pricing. Moreover, Doyle *et al.* [10] assess how the overconfidence affects decision making of grassroots employees in complex supply chain systems. Liu *et al.* [11] examine the effects of the dual-overconfident preferences and changing demand on decisions and utilities. They further develop a two-stage service procurement model in logistics service supply chain, and illustrate that a dynamic pricing mechanism eliminating the negative influence of overconfidence [12]. Li [5] clearly explain overconfidence bias can reduce the double marginalization effect. Furthermore, Xu *et al.* [13] analyse the impact of retailer's overconfidence on optimal pricing, ordering decisions in a duopolistic supply chain. However, the most existing papers only focus on the influence of overconfidence on decision variables, we extend the single-behavioural preference to dual-behavioural preferences in a supply chain system.

2.2 Fairness concern preference in supply chain

Fairness concern is an important social preference, which means that decision-makers not only pay attention to the maximum income of the individual, but also pay attention to whether the distribution of channel income or price is fair [14]. Cui *et al.* [15] incorporate fairness concern into a conventional dyadic supply chain consisting of one manufacturer and one retailer and considered its influence on channel coordination. Caliskan *et al.* [16] further extend the Cui's model to the nonlinear demand environment. Considering the different channel power, Li *et al.* [17] consider a retailer's fairness concern in a dual-channel supply chain model in which the dominant manufacturer has a direct channel, whereas the retailer is a follower. Nie *et al.* [6] further consider whether different contracts can coordinate supply chain performance when retailers show dual-fairness. However, Wang *et al.* [18] introduce the fairness concern of manufacturer into an e-commerce supply chain model. Pan *et al.* [19] explore the effect of fairness concern on decisions of supply chain consisting of a dominant retailer and two manufacturers. Different from the aforementioned papers, we develop a supply chain structure consisting of one rational manufacturer and one retailer with fairness concern and overconfidence.

Contrary to the earlier studies only considering the overconfidence or the fairness concern, we investigate the interactive effects of overconfidence and fairness concern on the supply chain performance. More recently, Zhang *et al.* [1] discuss optimal contract design in the joint effect of overconfidence and fairness concern on order quantity. In comparison, we mainly examine the interactive impacts of behavioural preferences on pricing and utilities in three scenarios, which gradually and clearly describes the interaction between overconfidence and fairness concern. In addition, they assume the market demand is uncertainty, but the market demand in this paper is deterministic to explore the effect of two behavioural preferences on decision variables and profit/utility by exactly mathematical formulas. It is interesting to find that utility of supply chain system may be beneficial to fairness concern when μ is in the range of $\left(\frac{\sqrt{2}-1}{2}, +\infty\right)$, and the wholesale price is not always increasing with overconfidence. More importantly, comparing to fairness concern, overconfidence is the most importantly behavioural preference affecting prices, sales effort and profit.

3. The model

3.1 Fundamental assumption and notation

A two-echelon supply chain where the manufacturer serving as the leader is unboundedly rational and the retailer serving as the follower has three conditions is presented; see Fig. 1. The manufacturer sells products to the retailer selling them to consumers. The manufacturer determines the wholesale price w . The retailer determines the retail price p and sales effort e according to the decision of manufacturer. We assume, without loss of generality, that transaction information between the manufacturer and the retailer is symmetrical, that the manufacturer's productivity can meet the retailer's ordering requirements and that the retailer can increase sales volume by improving the level of sales effort (e.g., advertising, giving gifts, etc.). Referring

to study from Karray [20] and Zeng et al. [21], let $c(e) = (1/2)\gamma e^2$, where $c(e)$ denotes the sales effort cost of the retailer and $\gamma (\gamma \geq 1)$ is the sale cost coefficient. The manufacturer and retailer are to maximize their expected profit or utility.

What we need to point out is that in Fig. 1, the single-behavioural preference refers to a retailer with overconfidence or fairness concern, while the two-behavioural preferences refer to a retailer who has both overconfident and fairness concern. We briefly present the notations and parameters used in this paper, as summarized in Table 1.

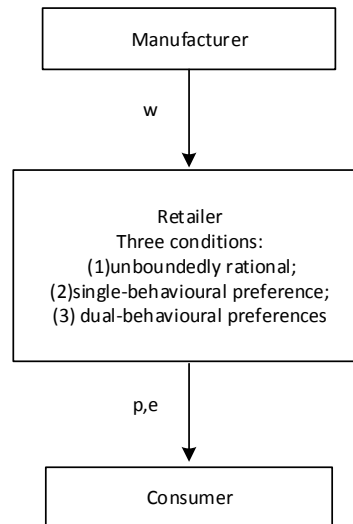


Fig. 1 The two-echelon supply chain structure

Table 1 Summary of parameters

Parameters	Description
a	The basic size of market demand ($a > 0$).
b	The price sensitivity coefficient of demand, i.e., the amount of marginal demand declines or rises when the price rises or falls by one unit ($b \geq 1$).
$c(e)$	The sales effort cost of the retailer, a monotonically increasing and concave function with e .
c_x	$x = m, r$ respectively denote the manufacturer's marginal cost of production and the retailer's marginal cost of selling, such as packaging, transportation, and maintenance fees.
D	The market demand function when the retailer is unboundedly rational.
D_o	The market demand function when the retailer has overconfidence, i.e., the overconfident retailer believes demand to be D_o in the market.
k	Retailer's sales effort output, i.e., the amount of marginal demand increases when the sales effort rises by one unit ($k \geq 1$).
γ	The sale cost coefficient ($\gamma \geq 1$)
β	The overconfidence parameter of the retailer, which represents the degree of overconfidence, $\beta \in (0, 1]$.
μ	The fairness concern parameter of the retailer, which represents the degree of fairness concern, $\mu > 0$.
π_i	$i = m, r, s$ denote the profit of supply chain members and supply chain system, respectively.
U_i	$i = r, s$ denote the utility function of the retailer and the two-echelon supply chain system, respectively.
j	The superscript $j = o, f, of$, respectively, indicates the scenarios of overconfidence, fairness concern and two-behavioural preferences of the retailer.
*	The superscript * indicates the equilibrium solution of variables or profits (utilities)

3.2 Basic model: Retailer with no behavioural preferences

In this section, a setting where there are no behavioural preferences between members is considered in a supply chain. In other words, both the manufacturer and the retailer are unboundedly rational and only interested in their own monetary payoffs. Next, the function of the market demand and the members' profits are built, after which the manufacturer plays a Stackelberg game with the retailer.

Based on the above problem description, the demand function in a two-echelon supply chain is defined as follows

$$D = a - bp + ke \quad (1)$$

The above price- and sales effort-sensitive demand function is a kind of linear model structure and has been systematically proven by many researchers (e.g., Mukhopadhyay *et al.* [22], Zheng *et al.* [23]). The advantages of the deterministic linear function are that it can give rise to explicit results for the optimal solutions in behavioural supply chain and is relatively easy to estimate parameters (i.e., k, b) in next numerical studies. In fact, the demand function is divided into two parts: $a - bp$ and ke . The meaning of these parameters (i.e., a, b and k) is shown in Table 1. When $e = 0$, D is equal to $(a - bp)$, which is the essential demand function and only related to p . We here assume $e > 0$ to study the effect of behavioural preferences on it. Thus, the market demand depends on retail price p as well as sales effort e , where is positively related to the demand but p is negatively correlated with it.

The profit of the manufacturer, the retailer and the whole channel is given below

$$\pi_m = (w - c_m)(a + ke - bp) \quad (2)$$

$$\pi_r = (p - w - c_r)(a + ke - bp) - \frac{1}{2}\gamma e^2 \quad (3)$$

$$\pi_s = (p - c_m - c_r)(a + ke - bp) - \frac{1}{2}\gamma e^2 \quad (4)$$

where the meaning of π_m, π_r, c_m, c_r and γ is described in Table 1. The π_s is the channel profit of the two-echelon supply chain, which is equivalent to the π_m plus π_r . Obviously, it is essential to impose additional inequality constraints on the parameters to guarantee the correct functions. Thus, these constraints are as follows

- $w > c_m, p > c_r$ and $w \leq p$.
- $(p - w - c_r) \geq 0$, there is a reason it may encourage the retailer to participate in order-sale activity; otherwise, if $(p - w - c_r) < 0$, the retailer cannot order products from the manufacturer, and thus the two-echelon supply chain does not exist [24].
- The π_r must be positive, which encourages the retailer to order more quantities from manufacturer and sell more goods to customers by increasing sales effort level.
- The D, π_m must be positive.

To obtain the subgame perfect equilibria for both manufacturer and retailer, the backward induction method is applied. The timeline of events is as follows. In stage 2 of the Stackelberg game, the retailer maximizes himself profit by choosing p and e for each product. According to the Hessian matrix $H(e, p) = \begin{pmatrix} -2b & 0 \\ 0 & -\gamma \end{pmatrix}$ in Eq. 3, we know that there exists unique optimal solution, i.e., e^*, p^* . In stage 1 of the game, according to the retailer's reaction, the manufacturer rationally chooses the optimal price based on c_m . Similarly, based on the first-order condition of π_m , we then obtain the optimal wholesale price w^* . The proof of equilibrium results is given in Appendix A. Lastly, we can derive

$$w^* = \frac{a + b(c_m - c_r)}{2b} \quad (5)$$

$$e^* = \frac{[a - b(c_m + c_r)]k}{2(2b\gamma - k^2)} \quad (6)$$

$$p^* = \frac{(b\gamma - k^2)[a + b(c_r + c_m)] + 2ab\gamma}{2b(2b\gamma - k^2)} \quad (7)$$

Substituting the above optimal decision variables Eqs. 5, 6 and 7 into Eqs. 2, 3 and 4, respectively, the equilibrium profit of the manufacturer, retailer, and supply chain system is obtained

$$\pi_m^* = \frac{[b(c_r + c_m) - a]^2\gamma}{4(2b\gamma - k^2)} \quad (8)$$

$$\pi_r^* = \frac{[b(c_r + c_m) - a]^2 \gamma}{8(2b\gamma - k^2)} \quad (9)$$

$$\pi_s^* = \frac{3\gamma[b(c_r + c_m) - a]}{8(2b\gamma - k^2)} \quad (10)$$

3.3 Behavioural model: Retailer with both overconfidence and fairness concern

In this section, the basic model is extended by incorporating overconfidence and fairness concern into the two-echelon supply chain. Specifically, we consider a setting where the manufacturer is unboundedly rational, whereas the retailer has both overconfidence and fairness concern in the two-echelon supply chain. Chen *et al.* [25] defined the retailer's overconfidence as an overestimation of the sales effort and market demand. What we convey in this paper is consistent with their idea of overconfidence. Thus, when the retailer is overconfident, the market demand is defined as $D_o = a + (k + \beta)e - bp$. The standard demand function can accurately measure the impact of overconfidence on demand. It is well known that the retailer's overconfidence parameter β is positively related to D_o , and the greater overconfidence degree β is, the greater is the effect of e on D_o . In Eq. 3, we replace D with D_o to obtain the utility function of overconfident retailer.

$$U_r^o = (p - w - c_r)[a + (k + \beta)e - bp] - \frac{1}{2}\gamma e^2 \quad (11)$$

The reference framework for fairness concern is the F-S model [26], and according to Zhang *et al.* [1], the retailer's utility with fairness concern is

$$U_r^f = \pi_r - \mu(\pi_m - \pi_r) = (1 + \mu)\pi_r - \mu\pi_m \quad (12)$$

wherein the μ is the degree of fairness concern, π_m and π_r is the profit of the rational retailer and manufacturer, respectively.

When the retailer with behavioural preferences, the rational manufacturer profit function is

$$\pi_m^{of} = (w - c_m)(a + ke - bp) \quad (13)$$

where Eq. 13 is equal to Eq. 2 because the manufacturer is unboundedly rational under the two scenarios.

According to Eqs. 12 and 13, the utility of retailer with the two behavioural preferences is as follows.

$$U_r^{of} = (1 + \mu)\{(p - w - c_r)[a + (k + \beta)e - bp] - \frac{1}{2}\gamma e^2\} - \mu(w - c_m)(a + ke - bp) \quad (14)$$

wherein when the retailer has both behavioural preferences, we can obtain Eq. 14 by replacing D with D_o in Eq. 12. Thus, the utility of the two-echelon supply chain system is equivalent to the manufacturer's profit plus the retailer's utility

The utility of the two-echelon supply chain is

$$U_s^{of} = (1 + \mu)\{(p - w - c_r)[a + (k + \beta)e - bp] - \frac{1}{2}\gamma e^2\} + (1 - \mu)(w - c_m)(a + ke - bp) \quad (15)$$

The Eqs. 14 and 15 can accurately measure the impact of behavioural preferences on utility, respectively.

In this section, the backward induction method is also applied to solve the subgame perfect equilibria for both the manufacturer and retailer. Following the same logic as in Section 3.2, we establish the Stackelberg game model in which the manufacturer maximizes profit and the retailer maximizes expected utility. As proof results in Appendix B, there are equilibrium solutions in the model when the retailer has the two-behavioural preferences.

$$w^{of*} = \frac{b^2 c_m \gamma (3\mu + 1) - b c_m k \beta (5\mu + 1) + (1 + \mu)[\beta(k + \beta)(b c_r - a) + b(a\gamma - b\gamma c_r - c_m \beta^2)]}{2b[b\gamma(2\mu + 1) - k\beta(3\mu + 1) - \beta^2(1 + \mu)]} \quad (16)$$

$$e^{of*} = \frac{[a - b(c_m + c_r)][2b\gamma\mu(k + 2\beta) - 2k\beta(2k + 3\beta) - \beta^3(2\mu + 1) + b\gamma(k + \beta)]}{2[2b\gamma - (k + \beta)^2][b\gamma(2\mu + 1) - k\beta(3\mu + 1) - \beta^2(1 + \mu)]} \quad (17)$$

$$p^{of*} = \frac{b\beta^3[\beta(1+\mu) + \gamma(1-\mu)] + [b\gamma^2(c_m + c_r)(\gamma - 1) + ab\gamma(3b\gamma - k^2) + k\beta(ak^2 + 3bc_m\beta^2)](1+2\mu)}{2b[2b\gamma - (k + \beta)^2][b\gamma(1+2\mu) - \beta^2(1+\mu) - k\beta(1+3\mu)]} + \frac{3bk\beta[(c_m + c_r)(b\gamma + k\beta) + c_r\beta^2](1+3\mu) + b\beta(c_m + c_r)(b\gamma\beta + k^3)(1+4\mu)}{2b[2b\gamma - (k + \beta)^2][b\gamma(1+2\mu) - \beta^2(1+\mu) - k\beta(1+3\mu)]} \quad (18)$$

Substituting the above optimal decision variables into Eqs. 13 and 14, respectively, the equilibrium profit of the manufacturer and the equilibrium utility of the retailer are obtained (due to the complexity of the utility function of the supply chain system, there is no display here, but it is equal to π_m^{of*} plus U_r^{of*}).

$$\pi_m^{of*} = \frac{(1+u)[b\gamma - (k + \beta)\beta]^2[a - b(c_m + c_r)]^2}{4b[2b\gamma - (k + \beta)^2][b\gamma(2\mu + 1) - k\beta(3\mu + 1) - \beta^2(1 + \mu)]^2} \quad (19)$$

$$U_r^{of*} = \frac{\alpha(1+u)[a - b(c_m + c_r)]^2}{8b[2b\gamma - (k + \beta)^2][1 + b\gamma(2\mu) - k\beta(1 + 3\mu) - \beta^2(1 + \mu)]^2} \quad (20)$$

wherein $\alpha = b\gamma(b^2\gamma^2 + k^2\beta^2)(1 + 2\mu)^2 + 2(b\beta\gamma)^2(3\mu^2 - \mu - 1) + 2\mu(k\beta)^3(1 + 4\mu) + 6\mu(k\beta^2)^2(1 + 3\mu) + 6k\mu\beta^5(1 + 2\mu) + 2\mu\beta^6(1 + \mu) - b\beta\gamma[2bk\gamma(1 + 2\mu)(1 + 3\mu) + 2k\beta^2(\mu^2 - 1)] + \beta^3(8\mu^2 + 4\mu + 1)$.

4. Analysis of equilibrium results

In this section, the effects of a retailer's behavioural preference on the decision variables and utility/ profit of members in the two-echelon supply chain are discussed. In fact, the model with the retailer's two-behavioural preferences can reduce back to the model with the retailer's single-behavioural preference, i.e., the retailer has only overconfidence or fairness concern. Therefore, for a clearer explanation, this section is divided into three parts: the retailer with overconfidence only, the retailer with fairness concern only and the retailer with behavioural preferences. Note that all propositions below are proven in Appendix C.

4.1 The retailer with overconfidence only

When the retailer's fairness concern parameter μ is equal to zero, the situation of a retailer with behavioural preferences degenerates into a retailer with overconfidence only. Thus, the following propositions are presented.

Proposition 1. When the retailer is overconfident, the wholesale price of the manufacturer is equivalent to the optimal wholesale price shown in the setting of the retailer's unbounded rationality, i.e., $w^{o*} = w^*$, while the retailer's optimal sales price p^{o*} and sales effort level e^{o*} are both positively correlated with β .

Proposition 1 states that when deciding the wholesale price, the rational manufacturer is not bothered by the retailer's overconfidence, whereas the level of sales effort and the retail price deviate from the optimal value of Section 3.2 where the retailer is rational. In addition, the retailer's overconfidence makes him overestimate his marketing capabilities, which directly motivate the retailer to make a greater sales effort (for example, increasing promotions). On the other hand, recalling Section 3.3, which discusses the theory of a retailer's overconfidence, the retailer's overconfidence parameter β is positively related to the market demand D_o , which means he overestimates the market demand. Thus, the superposition effect of the two aspects allows the retailer to continuously increase the retail price for a non-negative profit.

Proposition 2. When the retailer has overconfidence only, as its degree β increases, the profit of the manufacturer, the utility of the retailer, and the utility of the supply chain system also increase, i.e., π_m^{o*} , U_r^{o*} and U_s^{o*} are positively related to β .

Along with Proposition 1, Proposition 2 indicates that the overconfident retailer overestimates the market demand and sales effort level, and orders more products to meet customer demand as long as the wholesale price is constant, which can form a scale effect to increase profit for the manufacturer. The increase in the retailer's utility is attributed to two aspects: first,

overconfident behaviour makes him overestimate the market demand, and thus enhances his sales enthusiasm, which will in turn increase the sales volume of the product in general; second, the rise in the retail price increases the marginal revenue of every product. Thus, the increase in manufacturer profit and retailer utility ultimately leads to an increase in the utility of the supply chain system.

4.2 The retailer with fairness concern only

When the retailer's overconfident parameter β is equal to zero, the scenario of a retailer with behavioural preferences degenerates into the situation of a retailer with fairness concern only. Thus, the series of propositions below are obtained.

Proposition 3. When the retailer only pays attention to the fairness of income distribution in a two-echelon supply chain, the higher the degree of fairness concern μ , the lower is the wholesale price w^{f*} of the manufacturer, while the retail price p^{f*} and sales effort level e^{f*} are constant.

Proposition 3 shows in order to alleviate the retailer's unfair aversion, raise retailer's profitability and keep the competitiveness of the supply chain system, the manufacturer as the leader in the two-echelon supply chain needs to make concession, that is, he must lower the wholesale price, which means that the retailer can boost his bargaining power in the supply chain when he is concerned with the fairness of the relative revenue between the members. Nevertheless, the retailer's optimal decisions are not affected by his own fairness concern behaviour, the possible reason is that in order to maintain market share and improve demand, retailer does not adjust retail price and sales effort level because of fairness concern. The proposition is clearly different from the conclusions in other papers, e.g., Li *et al.* [17], Zheng *et al.* [27], where the retail price is both affected by fairness concern in a two-echelon supply chain.

Proposition 4. When the retailer has fairness concern only, the best utility of retailer U_r^{f*} grows with increasing degree of fairness concern μ , but the best profit of manufacturer π_m^{f*} falls with increasing μ . If μ is in the range of $(0, (\sqrt{2} - 1)/2)$, the U_s^{f*} decreases as μ increases; otherwise, if μ is in the range of $((\sqrt{2} - 1)/2, +\infty)$, the U_s^{f*} rises as μ increases.

Combining the change in wholesale price of Proposition 3, we clearly explain why the manufacturer's profit decreases with μ . In particular, Proposition 4 indicates that when μ is in the range of $(0, (\sqrt{2} - 1)/2)$, the bargaining power of the retailer is weak, which causes the decline in the manufacturer's profit margin to be higher than the increase in the retailer's utility margin; however, as μ increases, the retailer's bargaining power gradually increases, which causes the decline rate in the manufacturer's profit margin to be lower than the increase rate in the retailer's utility margin, and the supply chain system consists of a manufacturer and a retailer, thus its utility first declines and then rises. Moreover, from $\mu > 0$ and the conversion node of bargaining power for the retailer, i.e., $\mu = (\sqrt{2} - 1)/2$, it can be found that the retailer's bargaining space is very large.

4.3 The retailer with both overconfidence and fairness concern

When the retailer has both overconfidence and fairness concern, we obtain the following propositions. Due to the complex formula of the utility or profit between supply chain members, this section only analyses the effect of the two behavioural preferences (i.e., overconfidence and fairness concern) on decision variables.

Proposition 5. If other parameters remain unchanged, when the degree of overconfidence β is in the range of $((b\gamma - \sqrt{(b\gamma)^2 - 2b\gamma k})/2k, (b\gamma + \sqrt{(b\gamma)^2 - 2b\gamma k})/2k)$, the wholesale price decreases as β increases, and when β is outside the range, the wholesale price increases as β increases.

Unlike those in Proposition 1, Proposition 5 indicates that when the retailer has both overconfidence and fairness concern, if β is in the range of $((b\gamma - \sqrt{(b\gamma)^2 - 2b\gamma k})/2k, (b\gamma + \sqrt{(b\gamma)^2 - 2b\gamma k})/2k)$, the retailer's overconfidence will cause the manufacturer upstream of the supply chain to lower the wholesale price to encourage the retailer to order more products, which will ultimately increase the manufacturer's profit. When β is outside the range of $((b\gamma - \sqrt{(b\gamma)^2 - 2b\gamma k})/2k, (b\gamma + \sqrt{(b\gamma)^2 - 2b\gamma k})/2k)$, the rational manufacturer may realize

that the retailer's degree of overconfidence is unreasonable, and he can raise the wholesale price of the product to maximize his own revenue. This finding also coincides with the situation whereby merchants seek more profit in reality.

Proposition 6. If other parameters remain unchanged, the optimal wholesale price w^{of*} , the optimal retail price p^{of*} and the optimal sales effort e^{of*} are positively or negatively correlated with μ , which depends on the range of the degree of overconfidence β .

Proposition 6 is significantly different from Proposition 3, it shows that the correlation of the optimal decision variables (i.e., w^{of*} , p^{of*} , e^{of*}) and μ is determined by the range of β when other parameters are given; in other words, they are positively or negatively correlated with μ as the β changes, and β changes the decision results of members by first affecting the degree of μ in the supply chain. Thus, the above description illustrates that the degree of fairness concern changes according to the degree of overconfidence, and overconfidence is the main factor affecting the decision-making of the supply chain between the two behavioural preferences. Therefore, in actual operation management, when a leading manufacturer analyses or evaluates the effect of the two behavioural preferences of the retailer on business decisions, it is particularly important to pay attention to the retailer's overconfidence degree.

4.4 Comparative analysis of different models

For a clearer understanding of the influence of behavioural factors, optimal equilibrium solutions between basic model and behavioural model are analysed below. At the same time, due to the complexity of decision variables, we only compare the differences in wholesale price and in sales effort in two scenarios, respectively.

Proposition 7. The optimal wholesale price in Section 3.3 is less than that in Section 3.2, i.e. $\Delta w = w^{of*} - w^* < 0$. Nevertheless, the optimal sales effort in Section 3.3 is more than that in Section 3.2, i.e., $\Delta e = e^{of*} - e^* > 0$. In addition, both Δw and Δe increase with the degree of fairness concern μ .

Proposition 7 indicates that compared with the scenario of complete rationality, the two behavioural preferences are unfavourable to wholesale price, but they are conducive to the improvement of sales effort level. Moreover, the fairness concern degree μ can gradually expand the difference of decision variables (i.e., w , e) in different scenarios. Thus, we can show that the two-behavioural preferences are more favourable to retailer decision about sales effort.

5. Numerical analysis

We use numerical experiments to illustrate how the decision variables, profits and utilities change with behavioural preferences in the three scenarios. In this section, some parameter values are gathered by analysing the previous papers. According to [19] and [28], we set $\beta \in [0,1]$ and $\mu = \{0,1,2\}$ to illustrate the retailer's preference degree with regard to overconfidence and fairness concern in this paper. Moreover, we set $a = 100$, $\gamma = 1$, $c_m = 6$, $c_r = 8$, $b = 5$, and $k = 2$, which the constraint conditions of Stackelberg game models can be satisfied. Then, we can firstly obtain the equilibrium results $w^* = 9$, $p^* = 19.5$, $e^* = 5$, $\pi_m^* = 37.5$, $\pi_r^* = 18.75$, $\pi_s^* = 56.25$ in the first scenario.

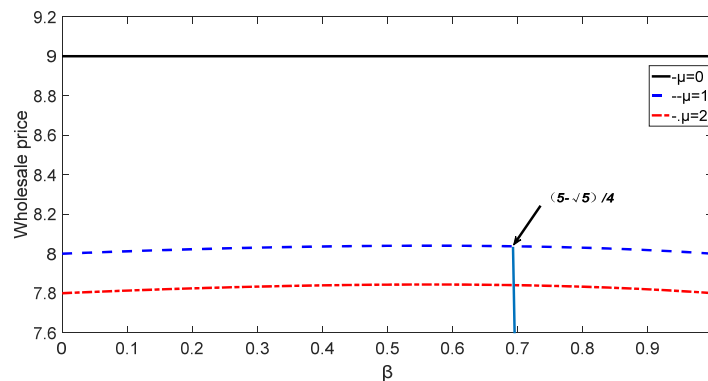


Fig. 2 Wholesale price under three scenarios

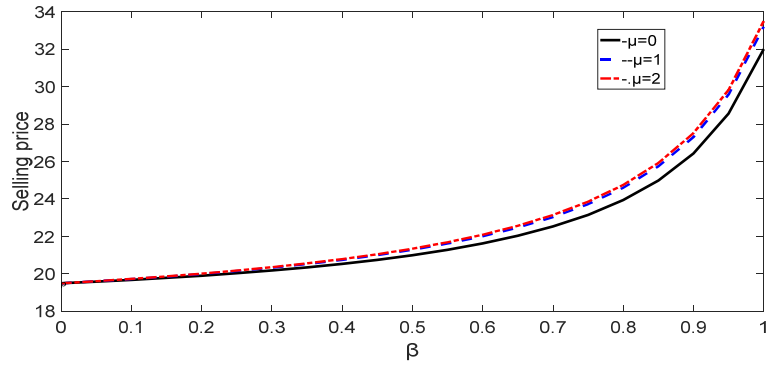


Fig. 3 Retail price under three scenarios

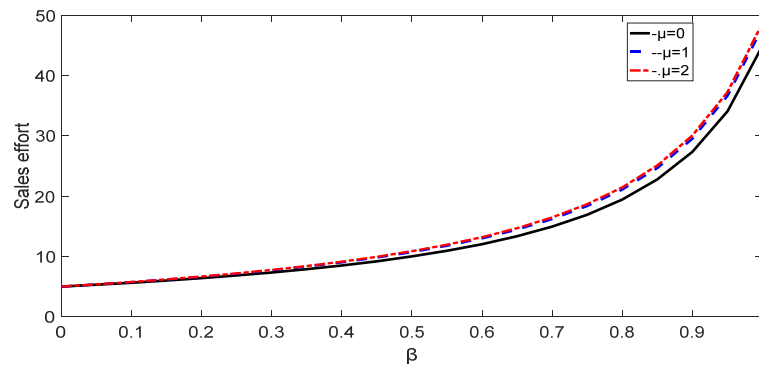


Fig. 4 Sales effort under three scenarios

When the retailer has single-behavioural preference or two-behavioural preferences, the changes in the decision variables, profit/utility and the utility of the two-echelon supply chain system are shown in Fig. 2 to Fig. 7.

In Fig. 2, when the retailer has overconfidence only ($\mu=0, 0 < \beta \leq 1$), the wholesale price of the manufacturer w^{0*} is equal to w^* that the retailer is rational ($\mu = 0, \beta = 0$); when the retailer has fairness concern only ($\beta = 0, \mu = \{1,2\}$), the wholesale price w^{f*} is negatively correlated with μ . This finding verifies the correctness of Propositions 1 and 3. Moreover, as β increases, the wholesale price in the two-behavioural preferences ($0 < \beta \leq 1, \mu = \{1,2\}$) slightly increases at the beginning and then decreases afterwards, the turning point being $(5 - \sqrt{5})/4$. That is, the w^{of*} is negatively correlated with β in the interval $[(5 - \sqrt{5})/4, 1]$, and it is positively correlated with β in the interval $(0, (5 - \sqrt{5})/4]$. This finding is consistent with the results of Proposition 5. In addition, the wholesale price in the dual-preferences scenario is a decreasing function of μ . In short, the wholesale price in the two-behavioural preferences is less than or equal to that in the single-behavioural preference. Therefore, the interaction of the two behavioural preferences is detrimental to the wholesale price.

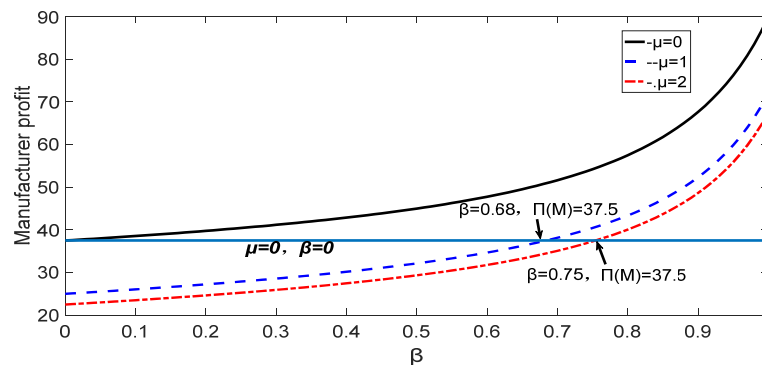


Fig. 5 Manufacturer profit under three scenarios

As shown in Figs. 3 and 4, both the sales price and sales effort are positively correlated with β and μ when retailer has confidence only ($\mu = 0, 0 < \beta \leq 1$) or dual-preferences ($0 < \beta \leq 1, \mu = \{1, 2\}$). Although the μ has a slight influence on sales price and sales effort, the two decisions are evidently affected by β . This result is mainly attributed to the fact that the retailer's overconfidence is the overestimation of market demand and sales effort, which prompts the retailer to raise the sale price and make a greater sales effort. In addition, the effects of μ on retail price and on sales effort under the scenario with fairness concern ($\beta = 0, \mu = \{1, 2\}$) and with rationality ($\beta = 0, \mu = 0$) are compared. Intuitively, the retailer concerning with the fairness of profit between members must pursue greater profits by improving price and sales effort level. However, as shown in Figs. 3 and 4, the retail price and sales effort remain unchanged. Therefore, this intuition is incorrect. Moreover, when retailer has overconfidence only, the values of both retail price and sales effort are more than these values when supply chain members are rational. In particular, for retailer, the gap of decision variables between the two scenarios is constantly expanding when β increases. In summary, based on the Fig. 3, Fig. 4 and analysis, we find that the retailer's overconfidence β is the main factor affecting retailer decision-making; in other words, behavioural preferences are beneficial to retailer's decisions, but are unfavourable to manufacturer's decision.

Fig. 5 illustrates that the optimal manufacturer's profit is an increasing function of β , but a decreasing function of μ in three scenarios. When the retailer is overconfident only, the manufacturer's profit π_m^{of*} is evidently larger than π_m^* , which confirms the result of Proposition 2. Besides, the difference of profit for manufacturer in the two scenarios (i.e., retailer has overconfidence only, retailer is rationality) quickly increases in β , but the difference of profit for manufacturer in the two scenarios (i.e., retailer has fairness concern only, retailer is rationality) gradually enlarges in μ . When the retailer has both overconfidence and fairness concern, the π_m^{of*} is lower than π_m^* at the beginning and then higher than π_m^* afterwards in the situation in which β and μ reach a certain value (i.e., $\beta = 0.68, \mu = 1$; $\beta = 0.75, \mu = 2$), respectively. In addition, in the three scenarios, the incremental slopes of the profit from the perspective of β is higher than the decrease in the slopes of profit from the perspective of μ , which also highlights that the main factor affecting the profit of the manufacturer is the retailer's overconfidence.

Both β and μ are beneficial to the retailer's utility in three scenarios in Fig. 6. Specially, as β increases, the utility growth rate of the retailer would be higher. The reasons for the above phenomena are the increase in the sales price and sales effort under the influence of the two behavioural preferences. At the same time, the gap of utility for retailer in the two scenarios (i.e., retailer has overconfidence only, retailer is rationality) gently grows in β , but the gap of utility for retailer in the two scenarios (i.e., the retailer has fairness concern only, the retailer is rationality) gradually decreases in μ . Comparing Fig. 6 with Fig. 5, we know that when retailer has overconfidence only the utility of retailer is less than or equal to the profit of manufacturer. However, when retailer has fairness concern only, the utility of retailer is gradually more than the profit of manufacturer. In particular, when retailer has both overconfidence and fairness concern, the utility of retailer is significantly larger than the profit of manufacturer, which may be due to the increase in retail price and sales of products. Therefore, the retailer's two-behavioural preferences are more favourable to his own utility than the manufacturers' profit.

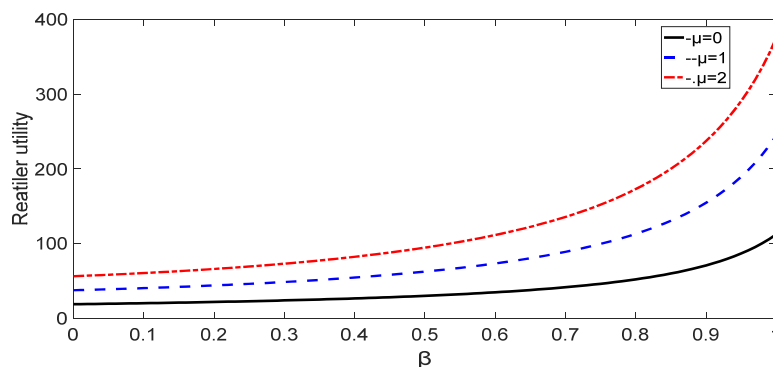


Fig. 6 Retailer's utility under three scenarios

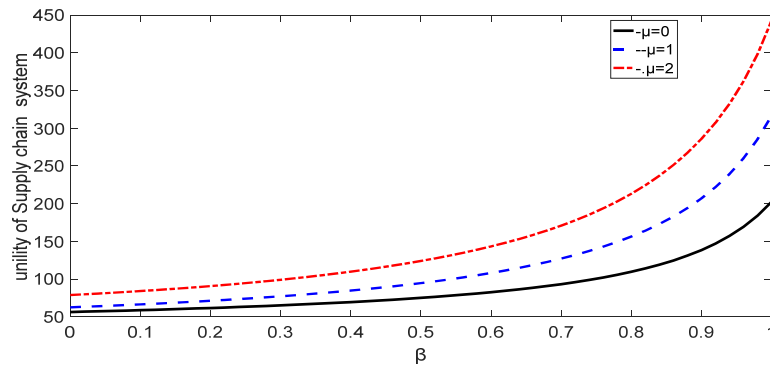


Fig. 7 Utility of the supply chain system under three scenarios

Based on Fig. 7, the utility of the supply chain system is positively correlated with β and μ in three scenarios. Combined with the analysis in Figs. 5 and 6, it is found that the change rate and trend of the utility curve between the retailer and the supply chain system are consistent. As we know, the utility of the supply chain system comprises the manufacturer's profit and the retailer's utility. Hence, the effect of the retailer's utility change on the utility of supply chain system is most obvious, which reflects to the importance of retailers' behavioural preferences in the two-echelon supply chain system.

6. Conclusion

In this paper, by establishing Stackelberg game models in which the manufacturer as a leader is unboundedly rational, and the retailer as a follower has two behavioural preferences, we analysed the effect of two behavioural preferences on equilibrium solutions, such as the wholesale price, sales price, sales effort and utility, under three scenarios and compared the equilibrium solutions with the results under the setting of the rational retailer. Based on these propositions and on the results from the numerical analysis, the conclusions and recommends of this paper are summarized as follows:

- When the retailer has a single behavioural preference, the overconfidence can improve the retailer's sales price and sales effort, while also increasing the utility (or profit) of the members and the supply chain system. Although fairness concern does not affect the retailer's sales price and sales effort level, they help improve the utility of the retailer and are not conducive to the manufacturer's pricing and profit, which ultimately cause fluctuation in the utility of the supply chain system. Moreover, the gap about retail price between the scenario of single-behavioural preference and the scenario of rationality is gradually increasing as overconfidence or as fairness concern. The same is true for sales effort.
- When the retailer has both overconfidence and fairness concern, the two behavioural preferences influence both the decision-making and utility (or profit) of the supply chain members. Between the two behavioural preferences, overconfidence is the main preference. Most importantly, the degree of fairness concern changes according to the degree of overconfidence, which can determine if fairness concern has a positive or negative influence on the wholesale price, sales price and sales effort. In addition, the positive effect of overconfidence on decision-making of the members in the supply chain is more pronounced than the retailer's fairness concern.

In practice, when information is transparently shared in the supply chain system, manufacturers, as the leader of the supply chain, can synthetically compare the historical transaction data of retailers with the actual market demand, and then evaluate the degree of the two behavioural preferences of retailers, especially the degree of overconfidence. For retailer, higher levels of overconfidence and fairness concern are needed to pursue more profits or utility. However, manufacturer should adopt reasonable incentives, e.g., contracts (Xiao *et al.* [3]), premium-

penalty mechanism, to guide retailers to control the degree of behaviour preferences, thereby promoting the improvement of profit and utility of the entire supply chain system.

One limitation of this study is that market demand is certain. As we know, the change in market environment and customer preference may induce a variety in market demand. Thus, we will build relevant models to study the effect of uncertainty and fuzziness base on different behavioural preferences, such as the uncertain demand function. In addition, combined with empirical economics, designing experimental studies in which the testers (retailers or manufacturers) are guided to control the degree of behavioural preferences within a reasonable range and reduce the negative effect is also on our research agenda. Finally, the profit distribution mechanism between the members will be designed and examined in the supply chain.

Acknowledgment

The research work presented in this article is supported by the National Natural Science Foundation of China (no. 71662011,71761015, 71940009).

References

- [1] Zhang, Z., Wang, P., Wan, M., Guo, J., Liu, J. (2020). Supply chain decisions and coordination under the combined effect of overconfidence and fairness concern, *Complexity*, Vol. 2020, Article ID 3056305, doi: [10.1155/2020/3056305](https://doi.org/10.1155/2020/3056305).
- [2] Lu, C.L. (2016). P&G's collaborative supply chain transformation, *Enterprise Management*, Vol. 10, 79-80.
- [3] Xiao, Q., Chen, L., Xie, M., Wang, C. (2020). Optimal contract design in sustainable supply chain: Interactive impacts of fairness concern and overconfidence, *Journal of the Operational Research Society*, 1-20, doi: [10.1080/01605682.2020.1727784](https://doi.org/10.1080/01605682.2020.1727784).
- [4] Yan, X., Chong, H.-Y., Zhou, J., Sheng, Z., Xu, F. (2020). Fairness preference based decision-making model for concession period in PPP projects, *Journal of Industrial & Management Optimization*, Vol. 16, No. 1, 11-23, doi: [10.3934/jimo.2018137](https://doi.org/10.3934/jimo.2018137).
- [5] Li, M. (2019). Overconfident distribution channels, *Production and Operations Management*, Vol. 28, No. 6, 1347-1365, doi: [10.1111/poms.12981](https://doi.org/10.1111/poms.12981).
- [6] Nie, T., Du, S. (2017). Dual-fairness supply chain with quantity discount contracts, *European Journal of Operational Research*, Vol. 258, No. 2, 491-500, doi: [10.1016/j.ejor.2016.08.051](https://doi.org/10.1016/j.ejor.2016.08.051).
- [7] Russo, J.E., Schoemaker, P.J.H. (1992). Managing overconfidence, *Sloan Management Review*, Vol. 33, No. 2, 7-17.
- [8] Ren, Y., Croson, D.C., Croson, R.T.A. (2017). The overconfident newsvendor, *Journal of the Operational Research Society*, Vol. 68, No. 5, 496-506, doi: [10.1057/s41274-016-0103-5](https://doi.org/10.1057/s41274-016-0103-5).
- [9] Kirshner, S.N., Shao, L. (2019). The overconfident and optimistic price-setting newsvendor, *European Journal of Operational Research*, Vol. 277, No. 1, 166-173, doi: [10.1016/j.ejor.2019.02.023](https://doi.org/10.1016/j.ejor.2019.02.023).
- [10] Doyle, J., Ojiako, U., Marshall, A., Dawson, I., Brito, M. (2020). The anchoring heuristic and overconfidence bias among frontline employees in supply chain organizations, *Production Planning & Control*, 1-18, doi: [10.1080/09537287.2020.1744042](https://doi.org/10.1080/09537287.2020.1744042).
- [11] Liu, B., Cai, G., Tsay, A.A. (2014). Advertising in asymmetric competing supply chains, *Production and Operations Management*, Vol. 23, No. 11, 1845-1858, doi: [10.1111/poms.12090](https://doi.org/10.1111/poms.12090).
- [12] Liu, W., Shen, X., Wang, D. (2018). The impacts of dual overconfidence behavior and demand updating on the decisions of port service supply chain: A real case study from China, *Annals of Operations Research*, Vol. 291, 565-604, doi: [10.1007/s10479-018-3095-5](https://doi.org/10.1007/s10479-018-3095-5).
- [13] Xu, L., Shi, X., Du, P., Govindan, K., Zhang, Z. (2019). Optimization on pricing and overconfidence problem in a duopolistic supply chain, *Computers & Operations Research*, Vol. 101, 162-172, doi: [10.1016/j.cor.2018.04.003](https://doi.org/10.1016/j.cor.2018.04.003).
- [14] Kahneman, D., Knetsch, J.K., Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market, *The American Economic Review*, Vol. 76, No. 4, 728-741.
- [15] Cui, T.H., Raju, J.S., Zhang, Z.J. (2007). Fairness and channel coordination, *Management Science*, Vol. 53, No. 8, 1303-1314, doi: [10.1287/mnsc.1060.0697](https://doi.org/10.1287/mnsc.1060.0697).
- [16] Caliskan-Demirag, O., Chen, Y., Li, J. (2010). Channel coordination under fairness concerns and nonlinear demand, *European Journal of Operational Research*, Vol. 207, No. 3, 1321-1326, doi: [10.1016/j.ejor.2010.07.017](https://doi.org/10.1016/j.ejor.2010.07.017).
- [17] Li, Q.-H., Li, B. (2016). Dual-channel supply chain equilibrium problems regarding retail services and fairness concerns, *Applied Mathematical Modelling*, Vol. 40, No. 15-16, 7349-7367, doi: [10.1016/j.apm.2016.03.010](https://doi.org/10.1016/j.apm.2016.03.010).
- [18] Wang, Y., Yu, Z., Shen, L. (2019). 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).
- [19] Pan, K., Cui, Z., Xing, A., Lu, Q. (2020). Impact of fairness concern on retailer-dominated supply chain, *Computers & Industrial Engineering*, Vol. 139, Article No. 106209, doi: [10.1016/j.cie.2019.106209](https://doi.org/10.1016/j.cie.2019.106209).
- [20] Karray, S. (2013). Periodicity of pricing and marketing efforts in a distribution channel, *European Journal of Operational Research*, Vol. 228, No. 3, 635-647, doi: [10.1016/j.ejor.2013.02.012](https://doi.org/10.1016/j.ejor.2013.02.012).

- [21] Zeng, L., Wang, J., Hu, Y. (2018). Retailer channel decisions of consumer electronics supply chain in a competitive environment, *Tehnički Vjesnik – Technical Gazette*, Vol. 25, No. 6, 1819-1828, doi: [10.17559/TV-20181101140915](https://doi.org/10.17559/TV-20181101140915).
- [22] Mukhopadhyay, S.K., Yao, D.-Q., Yue, X. (2008). Information sharing of value-adding retailer in a mixed channel hi-tech supply chain, *Journal of Business Research*, Vol. 61, No. 9, 950-958, doi: [10.1016/j.jbusres.2006.10.027](https://doi.org/10.1016/j.jbusres.2006.10.027).
- [23] Zheng, Z.L., Bao, X. (2019). The investment strategy and capacity portfolio optimization in the supply chain with spillover effect based on artificial fish swarm algorithm, *Advances in Production Engineering & Management*, Vol. 14, No. 2, 239-250, doi: [10.14743/apem2019.2.325](https://doi.org/10.14743/apem2019.2.325).
- [24] Jian, M., Wang, Y.L. (2018). Decision-making strategies in supply chain management with a waste-averse and stockout-averse manufacturer, *Advances in Production Engineering & Management*, Vol. 13, No. 3, 345-357, doi: [10.14743/apem2018.3.295](https://doi.org/10.14743/apem2018.3.295).
- [25] Chen, K.G., Song, X.F., Wang, X.Y., Huang, M. (2016). Joint pricing and production decisions with the overconfident sales agent, *Journal of Systems & Management*, Vol. 25, No. 3, 468-476.
- [26] Fehr, E., Schmidt, K.M. (1999). Theory of fairness, competition, and cooperation, *The Quarterly Journal of Economics*, Vol. 114, No. 3, 817-868, doi: [10.1162/003355399556151](https://doi.org/10.1162/003355399556151).
- [27] Zheng, X.-X., Li, D.-F., Liu, Z., Jia, F., Sheu, J.-B. (2019). Coordinating a closed-loop supply chain with fairness concerns through variable-weighted Shapley values, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 126, 227-253, doi: [10.1016/j.tre.2019.04.006](https://doi.org/10.1016/j.tre.2019.04.006).
- [28] Ancarani, A., Di Mauro, C., D'Urso, D. (2016). Measuring overconfidence in inventory management decisions, *Journal of Purchasing and Supply Management*, Vol. 22, No. 3, 171-180, doi: [10.1016/j.pursup.2016.05.001](https://doi.org/10.1016/j.pursup.2016.05.001).

Appendix A

When the members are unboundedly rational in the supply chain, their objectives are to maximize their respective gains. We apply the backward inductive method to model a Stackelberg game and obtain the equilibrium solutions.

In Eq. 3, the first-order and second-order derivative of π_r with respect to the decision variables p and e are the following

$$\begin{cases} \frac{\partial \pi_r}{\partial p} = a + ke - 2bp + b(w + c_r) = 0 \\ \frac{\partial \pi_r}{\partial e} = (p - w - c_r)k - \gamma e = 0 \end{cases} \quad (A1)$$

because $\frac{\partial^2 \pi_r}{\partial^2 p} = -2b < 0$, $\frac{\partial^2 \pi_r}{\partial^2 e} = -\gamma < 0$. Thus, the retailer has a unique equilibrium solution for retail price and sales effort. Thus, we obtain the following formula of the retail price and sales effort

$$\begin{cases} p = \frac{(b\gamma - k^2)(w + c_r) + a\gamma}{2b\gamma - k^2} \\ e = \frac{[a - b(w + c_r)]k}{2b\gamma - k^2} \end{cases} \quad (A2)$$

Based on the demand function, i.e., $D = a + ke - bp$, then substituting Eq. A2 into D , we can obtain $D = \frac{b\gamma[a - b(c_m + c_r)]}{2b\gamma - k^2}$. Recall that the assumption $D > 0$ holds. Therefore, $2b\gamma - k^2 > 0$, $a - b(c_m + c_r) > 0$.

By substituting Eq. A2 into Eq. 2, we obtain the function $\pi_m(w)$.

$$\pi_m(w) = \frac{b\gamma(w - c_r)[a - b(w + c_r)]}{2b\gamma - k^2} \quad (A3)$$

In Eq. A3, the first-order and second-order conditions of π_m with respect to the decision variable w are as follows: $\frac{d\pi_m}{dw} = \frac{b\gamma[a - b(2w + c_r - c_m)]}{2b\gamma - k^2} = 0$, $\frac{d^2\pi_m}{d^2w} = \frac{-2\gamma b^2}{2b\gamma - k^2} < 0$. Thus, the manufacturer has a unique equilibrium solution regarding wholesale price. The equilibrium solution of the wholesale price is obtained.

$$w^* = \frac{a + b(c_m - c_r)}{2b} \quad (A4)$$

Substituting Eq. A4 into Eq. A2, we obtain the optimal solution regarding retail price and sales effort.

$$\begin{cases} p^* = \frac{(b\gamma - k^2)[a + b(c_r + c_m)] + 2ab\gamma}{2b(2b\gamma - k^2)} \\ e^* = \frac{[a - b(c_m + c_r)]k}{2(2b\gamma - k^2)} \end{cases} \quad (A5)$$

Lastly, we substitute these equilibrium solutions of w^* , e^* and p^* into Eq. 2, Eq. 3 and Eq. 4, and the equilibrium profits of the manufacturer, retailer, and supply chain system are obtained, as shown in Eqs.8, 9, and 10.

Appendix B

First, must to prove that the retailer utility function U_r^{of} has a unique equilibrium solution. The first-order and second-order derivative of U_r^{of} with respect to p and e are as follows

$$\begin{cases} \frac{\partial U_r^{of}}{\partial p} = (1 + \mu)[a + (k + \beta)e - bp - (p - w - c_r)] - \mu(w - c_m)(a - bp + ke) = 0 \\ \frac{\partial U_r^{of}}{\partial e} = (1 + \mu)[(p - w - c_r)(k + \beta) - \gamma e] - \mu k(w - c_m) = 0 \end{cases} \quad (B1)$$

because $\frac{\partial^2 U_r^{of}}{\partial^2 p} = -2b(1 + \mu) < 0$, $\frac{\partial^2 U_r^{of}}{\partial^2 e} = -(1 + \mu)\gamma < 0$. Thus, the retailer with both overconfidence and fairness concern has a unique equilibrium solution regarding retail price and sales effort. Thus, we obtain the formula of the retail price and sales effort

$$\begin{cases} p^{of} = \frac{-(1 + \mu)(k + \beta)^2(w + c_r) + a\gamma(1 + \mu) + b\gamma w(1 + 2\mu) + k\mu(k + \beta)(c_m - w) + b\gamma[(1 + \mu)c_r - \mu c_m]}{(1 + \mu)[2b\gamma - (k + \beta)^2]} \\ e^{of} = \frac{(k + \beta)[(1 + \mu)\mu - b(w + c_r) + b\mu(c_m - c_r)] + (a - 2bw)k\mu}{(1 + \mu)[2b\gamma - (k + \beta)^2]} \end{cases} \quad (B2)$$

Substituting Eq. B2 into $D_o = a - bp + (k + \beta)e$, we can obtain

$D_o = \frac{b[\beta\mu(k + \beta)(w - c_m) + (1 + \mu)\gamma(a - bc_r) + b\gamma(w(1 + 2\mu) - \mu c_m)]}{(1 + \mu)[2b\gamma - (k + \beta)^2]}$. Recall the assumption $D_o > 0$. In addition, the additional inequality constraints in Section 3.2, $w > c_m$, $p > c_r$, $a - bc_r > 0$. Therefore, $[\beta\mu(k + \beta)(w - c_m) + (1 + \mu)\gamma(a - bc_r) + b\gamma(w(1 + 2\mu) - \mu c_m)] > 0$ and $2b\gamma > (k + \beta)^2$, i.e., $0 \leq \beta \leq \sqrt{2b\gamma} - k$. In addition, substituting Eq. B2 into Eq. 14, we obtain $\pi_m^{of}(w)$.

$$\pi_m^{of}(w) = \frac{(w - c_r)\{\beta(k + \beta)[bc_r\mu - a(1 + \mu)] + b\beta w[(1 + 3\mu)k + \beta(1 + \mu)] - b^2\gamma[(1 + \mu)c_r + w(1 + 2\mu)] + bc_m\mu(b\gamma - 2k\beta)\}}{(1 + \mu)[2b\gamma - (k + \beta)^2]} \quad (B3)$$

In Eq. B3, the first-order and second-order conditions of $\pi_m^{of}(w)$ with respect to w is $\frac{d\pi_m}{dw} = 0$, $\frac{d^2\pi_m}{d^2w} = \frac{-2b\{[2b\gamma - (k + \beta)^2](1 + \mu) + k(k + \beta) + \mu k(k - \beta)\}}{(1 + \mu)[2b\gamma - (k + \beta)^2]} < 0$, respectively. Thus, the manufacturer has a unique equilibrium solution regarding wholesale price when the retailer has both overconfidence and fairness concern. The equilibrium solution is

$$w^{of*} = \frac{b^2c_m\gamma(3\mu + 1) - bc_mk\beta(5\mu + 1) + (1 + \mu)[\beta(k + \beta)(bc_r - a) + b(a\gamma - b\gamma c_r - c_m\beta^2)]}{2b[b\gamma(2\mu + 1) - k\beta(3\mu + 1) - \beta^2(1 + \mu)]} \quad (B4)$$

Substituting Eq. B4 into Eq. B2, the optimal solutions regarding retail price and sales effort are obtained.

$$\begin{cases} e^{of*} = \frac{[a - b(c_m + c_r)][2b\gamma\mu(k + 2\beta) - 2k\beta(2k + 3\beta) - \beta^3(2\mu + 1) + b\gamma(k + \beta)]}{2[2b\gamma - (k + \beta)^2][b\gamma(2\mu + 1) - k\beta(3\mu + 1) - \beta^2(1 + \mu)]} \\ p^{of*} = \frac{b\beta^4(1 + \mu) + b\gamma\beta(1 - \mu) + [b\gamma^2(c_m + c_r)(\gamma - 1) + ab\gamma(3b\gamma - k^2) + k\beta(ak^2 + 3bc_m\beta^2)](1 + 2\mu) + 6abk\gamma\beta\mu}{2b[2b\gamma - (k + \beta)^2][b\gamma(1 + 2\mu) - \beta^2(1 + \mu) - k\beta(1 + 3\mu)]} + \frac{3bk\beta[(c_m + c_r)(b\gamma + k\beta) + c_r\beta^2](1 + 3\mu) + b\beta(c_m + c_r)(b\gamma\beta + k^3)(1 + 4\mu) + ak\beta^2(k\mu + \beta)(3 + 4\mu) - ab\gamma\beta(4 + 5\mu)(\beta - 7k)}{2b[2b\gamma - (k + \beta)^2][b\gamma(1 + 2\mu) - \beta^2(1 + \mu) - k\beta(1 + 3\mu)]} \end{cases} \quad (B5)$$

Lastly, substituting the equilibrium solutions into Eqs.13 and 14, the equilibrium profit of the manufacturer and retailer are obtained, as shown in Eqs.19 and 20.

Appendix C

Proof of Proposition 1

When the retailer has overconfidence only, we set μ to 0 in Eqs.(16), (17) and (18), and the optimal wholesale price, retail price and sales effort are obtained. It is easy to see that

$$w^{o*} = \frac{a+b(c_m-c_r)}{2b} \quad (C1)$$

$$p^{o*} = \frac{[b\gamma-(k+\beta)^2][a+b(c_r+c_m)]+2ab\gamma}{2b[2b\gamma-(k+\beta)^2]} \quad (C2)$$

$$e^{o*} = \frac{(k+\beta)[a-b(c_m+c_r)]}{2[2b\gamma-(k+\beta)^2]} \quad (C3)$$

Next, the first-order conditions of w^{o*} , p^{o*} , e^{o*} with respect to β are as follows. $\Delta w = w^{o*} - w^* = 0$, $\frac{dp^{o*}}{d\beta} = \frac{(k+\beta)\gamma[a-b(c_m+c_r)]}{[2b\gamma-(k+\beta)^2]^2}$, $\frac{de^{o*}}{d\beta} = \frac{[a-b(c_m+c_r)][2b\gamma+(k+\beta)^2]}{2[2b\gamma-(k+\beta)^2]^2}$. According to Appendix A and B, since $a - b(c_m + c_r) > 0$, $2b\gamma - (k + \beta)^2 > 0$, hence $\frac{dp^{o*}}{d\beta} > 0$, $\frac{de^{o*}}{d\beta} > 0$.

Proof of Proposition 2

Similar to the proof process for Proposition 1, let $\mu = 0$ in Eqs.19 and 20, the manufacturer's profit, the retailer's utility and the utility of the supply chain when the retailer has overconfidence only are obtained.

$$\pi_m^{o*} = \frac{[a-b(c_m+c_r)]^2[br-\beta(k+\beta)]}{4b[2b\gamma-(k+\beta)^2]} \quad (C4)$$

$$U_r^{o*} = \frac{\gamma[a-b(c_m+c_r)]^2}{8[2b\gamma-(k+\beta)^2]} \quad (C5)$$

$$U_s^{o*} = \pi_m^{o*} + U_r^{o*} = \frac{[a-b(c_m+c_r)]^2[3b\gamma-2\beta(k+\beta)]}{8b[2b\gamma-(k+\beta)^2]} \quad (C6)$$

Thus, $\frac{d\pi_m^{o*}}{d\beta} = -\frac{[a-b(c_m+c_r)]^2[2b\gamma\beta-k(k+\beta)^2]}{4b[2b\gamma-(k+\beta)^2]^2}$, $\frac{dU_r^{o*}}{d\beta} = \frac{(k+\beta)\gamma[a-b(c_m+c_r)]^2}{4[2b\gamma-(k+\beta)^2]^2}$, respectively. According to Appendix B, since $2b\gamma - (k + \beta)^2 > 0$, hence, $\frac{d\pi_m^{o*}}{d\beta} > 0$, $\frac{dU_r^{o*}}{d\beta} > 0$. Meanwhile, $U_s^{o*} = \pi_m^{o*} + U_r^{o*}$, that is, manufacturer profit and retailer utility together affect the change of utility of the supply chain system. Thus, the U_s^{o*} has same property, i.e., $\frac{dU_s^{o*}}{d\beta} > 0$.

Proof of Proposition 3

When the retailer has fairness concern only, i.e., $\beta = 0$, $\mu > 0$. Thus, let $\beta = 0$ in Eqs.16, 17 and 18, and the optimal wholesale price, retail price and sales effort for each member in the scenarios are as follows:

$$w^{f*} = \frac{(1+\mu)(a-bc_r)+(3\mu+1)bc_m}{2b(2\mu+1)} \quad (C7)$$

$$p^{f*} = \frac{(b\gamma-k^2)[a+b(c_r+c_m)]+2ab\gamma}{2b(2b\gamma-k^2)} \quad (C8)$$

$$e^{f*} = \frac{[a-b(c_r+c_m)]k}{2(2b\gamma-k^2)} \quad (C9)$$

Then, the first-order conditions of w^{f*} , p^{f*} , e^{f*} with respect to μ is $\frac{dw^{f*}}{d\mu} = -\frac{a-b(c_m+c_r)}{2b(2\mu+1)^2}$, $p^{f*} = p^*$, $e^{f*} = e^*$. According to Appendix A, $a - b(c_m + c_r) > 0$. Hence, $\frac{dw^{f*}}{d\mu} < 0$. Interestingly, comparing p^{f*} and e^{f*} with p^* and e^* , the optimal price and sales effort of the retailer with fairness concern are equal to the equilibrium results of the retailer without fairness concern.

Proof of Proposition 4

Similar to the proof process for Proposition 3, let $\beta = 0$ in Eqs.19 and 20, the equilibrium manufacturer's profit, the retailer's utility and utility of supply chain when the retailer has fairness concern only are obtained.

$$\pi_m^{f*} = \frac{(1+\mu)\gamma[a-b(c_m+c_r)]^2}{4(1+2\mu)(2b\gamma-k^2)} \tag{C10}$$

$$U_r^{f*} = \frac{(1+\mu)\gamma[a-b(c_m+c_r)]^2}{8(2b\gamma-k^2)} \tag{C11}$$

$$U_s^{f*} = \pi_m^{f*} + U_r^{f*} = \frac{(1+\mu)(2\mu+3)\gamma[a-b(c_m+c_r)]^2}{8(2\mu+1)(2b\gamma-k^2)} \tag{C12}$$

and $\frac{d\pi_m^{f*}}{d\mu} = -\frac{\gamma[a-b(c_m+c_r)]^2}{4(1+2\mu)^2(2b\gamma-k^2)} < 0$; $\frac{dU_r^{f*}}{d\mu} = \frac{\gamma[a-b(c_m+c_r)]^2}{8(2b\gamma-k^2)} > 0$; $\frac{dU_s^{f*}}{d\mu} = \frac{[a-b(c_m+c_r)]^2(4\mu^2+4\mu-1)\gamma}{8(2\mu+1)^2(2b\gamma-k^2)}$.

According to Appendix A, $2b\gamma - k^2 > 0$, thus, $\frac{d\pi_m^{f*}}{d\mu} > 0$ and $\frac{dU_r^{f*}}{d\mu} > 0$. For the function $\frac{dU_s^{f*}}{d\mu}$, let $(4\mu^2 + 4\mu - 1) = 0$, $\mu_1 = \frac{1-\sqrt{2}}{2}$ and $\mu_2 = \frac{-1+\sqrt{2}}{2}$ are obtained. Since $\mu > 0$, thus, μ_1 is omitted; Given the $\mu \in (0, \mu_2)$, the $(4\mu^2 + 4\mu - 1)$ is less than zero by calculating; then the $\frac{dU(\pi^f)}{d\mu} < 0$ can be observed. Whereas given the $\mu \in (\mu_2, +\infty)$, we can receive the $\frac{d\pi^f}{d\mu} > 0$.

Proof of Proposition 5

When the retailer has dual-preferences, the first-order partial derivative of w^{of*} with respect to β is as follows:

$$\frac{\partial w^{of*}}{\partial \beta} = \frac{\mu(1+\mu)[a-b(c_m+c_r)][b\gamma+2\beta(k\beta-b\gamma)]}{2b[b\gamma(1+2\mu)-k\beta(1+3\mu)-\beta^2(1+\mu)]^2} \tag{C13}$$

According to Appendix A, $a - b(c_m + c_r) > 0$, and $b > 0, \mu > 0, \mu(1 + \mu) > 0$, and the denominator of $\frac{\partial w^{of*}}{\partial \beta}$ is positive. Thus, it is easy to see that $\frac{\partial w^{of*}}{\partial \beta}$ is less than zero or more than zero depending on the degree of $b\gamma + 2\beta(k\beta - b\gamma)$.

Let $b\gamma + 2\beta(k\beta - b\gamma) = 0$, we can easily obtain the $\beta_1 = \frac{b\gamma - \sqrt{(b\gamma)^2 - 2b\gamma k}}{2k}$ and $\beta_2 = \frac{b\gamma + \sqrt{(b\gamma)^2 - 2b\gamma k}}{2k}$. According to Appendix B, $2b\gamma > (k + \beta)^2$, and β is in the interval $(0,1]$. Thus, we obtain $(b\gamma)^2 - 2b\gamma k \in (0, b\gamma)$ and $0 < \beta_1 < \beta_2 < 2k$. Therefore, $\frac{\partial w^{of*}}{\partial \beta} > 0$ only when $\beta \in (0, \beta_1)$ or $\beta \in (\beta_2, 1]$; $\frac{\partial w^{of*}}{\partial \beta} < 0$ only when $\beta \in (\beta_1, \beta_2)$. Note that the partial derivatives of p^{of*}, e^{of*} with respect to β are more complex. There is no analysis listed here, but their analysis will be carried out through numerical analysis.

Proof of Proposition 6

When the retailer has dual-preferences, the first-order derivative of w^{of*}, e^{of*} and p^{of*} with respect to μ are as follows:

$$\frac{\partial w^{of*}}{\partial \mu} = \frac{(b\gamma-2k\beta)[br-\beta(k+\beta)][a-b(c_m+c_r)]}{2b[b\gamma(2\mu+1)-k\beta(3\mu+1)-\beta^2(1+\mu)]^2} \tag{C14}$$

$$\frac{\partial e^{of*}}{\partial \mu} = \frac{[a-b(c_m+c_r)][b\gamma-\beta(k+\beta)]}{2[b\gamma(2\mu+1)-k\beta(3\mu+1)-\beta^2(1+\mu)]^2} \tag{C15}$$

$$\frac{\partial p^{of*}}{\partial \mu} = \frac{k\beta[br-\beta(k+\beta)][a-b(c_m+c_r)]}{2b[b\gamma(1+2\mu)-(1+\mu)\beta^2-(1+3\mu)k\beta]^2} \tag{C16}$$

Since $(b\gamma - 2k\beta)$ may be greater than zero or less than zero, we assume $2k\beta > b\gamma$ or $2k\beta < b\gamma$. Let $b\gamma - \beta(k + \beta) = 0$; it is easily to obtain $\beta_3 = \frac{k-\sqrt{k^2+4b}}{-2}$ and $\beta_4 = \frac{k+\sqrt{k^2+4b}}{-2}$. Recall that $\beta \in (0,1]$, and $0 < k < \sqrt{k^2+4b} < 2$ can be easily observed. Hence, the β_4 is omitted, and β_3 is reserved. The β_3 has two intervals $(0,1)$ or $[1, +\infty)$. Here, we only analyse the interval $(0,1)$,

because it is easily observed that the relationship between optimal wholesale price and μ when β_3 is in the range of $[1, +\infty)$.

We first present the relationship between w^{of*} and μ when $2k\beta > b\gamma$. It can be easily verified that the denominator of $\frac{\partial w^{of*}}{\partial \mu}$ is positive and that $a - b(c_m + c_r)$ and b are positive. Therefore, $\frac{\partial w^{of*}}{\partial \mu} < 0$ only when $\beta \in (0, \beta_3)$, and $\frac{\partial w^{of*}}{\partial \mu} > 0$ only when $\beta \in (\beta_3, 1]$. Next, we present the relationship between w^{of*} and μ when $2k\beta < b\gamma$; in the same condition, we obtain where $\frac{\partial w^{of*}}{\partial \mu} > 0$ only when $\beta \in (0, \beta_3)$, and $\frac{\partial w^{of*}}{\partial \mu} < 0$ only when $\beta \in (\beta_3, 1]$. The proofs for $\frac{\partial e^{of*}}{\partial \mu} > 0, \frac{\partial p^{of*}}{\partial \mu} > 0$ or $\frac{\partial e^{of*}}{\partial \mu} < 0, \frac{\partial p^{of*}}{\partial \mu} < 0$ are similar.

Proof of Proposition 7

Comparing the optimal wholesale price and sales effort under basic model and behavioural model, we get the solutions:

$$\Delta w = w^{of*} - w^* = -\frac{(a-b(c_m+c_r))(2k\beta+b\gamma)\mu}{2b(b\gamma(1+2\mu)+\beta((1+\mu)\beta+(1+3\mu)k))} \tag{C17}$$

$$\Delta e = e^{of*} - e^* = \frac{k(a-b(c_m+c_r))(k+2\beta)}{2(2b\gamma-k^2)(b\gamma(1+2\mu)-k\beta(1+3\mu)-\beta^2(1+\mu))} \tag{C18}$$

Since $\beta \in (0,1], a - b(c_m + c_r) > 0$ and $2br - (k + \beta)^2 > 0$, in addition, the optimal sales effort must be greater than 0, thus $(b\gamma(1 + 2\mu) > k\beta(1 + 3\mu) + \beta^2(1 + \mu))$. Therefore $\Delta w < 0$ and $\Delta e > 0$.

$$\frac{\partial \Delta w}{\partial \mu} = \frac{\beta(a-b(c_m+c_r))(2k\beta+b\gamma)(\beta+k)}{2b(b\gamma(1+2\mu)+\beta((1+\mu)\beta+(1+3\mu)k))^2} \tag{C19}$$

$$\frac{\partial \Delta e}{\partial \mu} = \frac{k(a-b(c_m+c_r))(k+2\beta)(2b\gamma-\beta^2-3k\beta)}{4(2b\gamma-k^2)(b\gamma(1+2\mu)-\beta^2(1+\mu)-\beta k(1+3\mu))^2} \tag{C20}$$

According to $a - b(c_m + c_r) > 0$, obviously, $\frac{\partial \Delta w}{\partial \mu} > 0$. In addition, since $\beta \in (0,1], b \geq 1, \gamma \geq 1, 2b\gamma > \beta^2 + 3k\beta$, and $2b\gamma > \beta^2 + 3k\beta$, thus, $\frac{\partial \Delta e}{\partial \mu} > 0$.

Computational analysis of cavitation at the tongue of the volute of a centrifugal pump at overload conditions

Hu, Q.^a, Yang, Y.^a, Cao, W.^{a,*}

^aResearch Institute of Fluid Engineering Equipment Technology, Jiangsu University, P.R. China

ABSTRACT

Volute pump is the most common used centrifugal pump. As cavitation highly contributes to deteriorating the performance of the pump, anti-cavitation performance is one of its design specifications. To clarify the cavitation evolution at the tongue of the volute of a centrifugal pump at overload conditions and its influence on the flow field in the impeller, numerical simulation with ANSYS CFD and a hydraulic test were conducted on a volute pump at several flow rates above optimal value. The cavity distribution and the blade loading distribution were analyzed. And the characteristics of the pressure fluctuation of the monitoring points located in volute casing were obtained and compared with each other. Results showed that cavitation may first emerge at the tongue rather than the impeller inlet at overload conditions. The alternative stress resulting from periodical radial force increases obviously as the extent of cavitation at the tongue. Meanwhile, the mean cavity length grows as each blade comes close to the tongue, and causes a decrease in performance because of a reduction or closure of flow passages. The pressure pulsation in the volute is consistent with the blade passing frequency whether cavitation occurs or not, while the pulsation intensity increases obviously after cavitation inception. From the first section to the eighth section of volute, the pulsation intensity of impeller outlet decreases gradually. The results are then compared to provide a reference for the optimum design of the anti-cavitation performance of centrifugal pump.

© 2020 CPE, University of Maribor. All rights reserved.

ARTICLE INFO

Keywords:
Centrifugal pump;
Numerical simulation;
Computational fluid dynamics (CFD);
Tongue;
Cavitation;
Blade loading;
Pressure fluctuation

**Corresponding author:*
[cwg@ujs.edu.cn](mailto:cwd@ujs.edu.cn)
(Cao, W.)

Article history:
Received 18 August 2020
Revised 14 September 2020
Accepted 16 September 2020

1. Introduction

The generation of cavitation can cause severe damage in hydraulic machinery. In liquid flows, the vaporization of liquid is generally due to the local pressure becomes lower than the saturated vapor pressure [1]. Volute pumps are among the most used centrifugal pump types, mainly in mining, petrochemical, and paper industries. Cavitation is a major undesirable phenomenon, which occurs in the operation of centrifugal pumps [2]. The occurrence of cavitation in passages affects the energy change between the wall surface and liquid, which leads to the breakdown of performance, surface erosion, and even the equipment failure [3, 4]. In case of centrifugal pumps, the cavitation inception generally takes place on the suction surface of the blades near their leading edge, and the cavity presents an asymmetrical feature, which is caused by the interaction between the blade and tongue of volute.

On this basis, cavitation in centrifugal pumps have been extensively studied [5-12]. Most studies mainly focused on the cavitation detection in impellers and performance improvement through the application of cavitating flow simulation. However, only a few studies on the cavitating

tion at the tongue of volute have been conducted based on visualization experiment and simulation computation. At overload conditions, flow separation, characterized by unsteady vortex shedding from the tongue, causes the low pressure zone near the tongue. The present study deals with a special case, where cavitation near the tongue of the volute occurs prior to the cavitation at the leading edge of blades. The head drop curve also has a knee shape that head remain constant with the decrease of *NPSH* and rapidly decreases at critical point. In this case the breakdown of performance due to cavitation at the tongue rather than cavitation in impeller. With pressure fluctuation and visualization experiment, and combining numerical simulation, the cavitation structure at the tongue and its influence to the flow field in impeller was investigated. The results were then compared to provide a reference for the optimum design of centrifugal pump.

2. State of the art

The investigation of flow field in the vicinity of or inside cavitation is complex in spite of remarkable progress on the experimental technique and numerical calculation. R.F. Kunz *et al.* [13] predicted the occurrence and evolution of single airfoil cavitation by using a two-phase flow model based on the Navier-Stokes equation. On this basis, the present study conducted a numerical simulation analysis and research on a centrifugal pump and investigated the anti-cavitation performance. Luo *et al.* [14] compared the cavitation prediction for a centrifugal pump with and without volute casing. Both models predicted the performance deterioration caused by cavitation, while the asymmetrical feature of cavitating flow exists when the calculation domain with volute casing is applied. And the performance deterioration caused by the asymmetrical cavitation is overestimate based on the experiment data. Rudolf *et al.* [15] provide high-speed photographic observation of successive stages of unsteady cavitation at the tongue of the volute of a centrifugal pump. The experiments were carried out at a flow rate above optimal value and at 3 % head drop conditions. The cloud cavitation at the tongue is similar to the one on single hydrofoils. However, only direct visualization was conducted to explain the evolution of cavitation structures at the tongue of centrifugal pump at overload conditions, numerical method was not employed to investigate the cavitation phenomenon at the tongue. To solve the cavitation damage problem near the casing tongue of an afterburner fuel pump, Xue *et al.* [16] investigated the separated flow characteristics and pressure fluctuations around the tongue by large eddy simulations. It is found that at low flow conditions, the fuel backflow from diffuser to the annular chamber triggers serious separation, which lowers the local static pressure to less than zero. The location of the separated vortex is in accord with the cavitation damage core region. Meng *et al.* [17] simulated and analyzed the complex transient cavitation flow patterns inside a centrifugal pump. The cavities in the passages exhibit an obvious life cycle with a frequency corresponding to the impeller rotation frequency under off-design conditions. The asymmetric cavitation is associated with the uneven pressure distribution on volute and impeller-tongue interaction. Dular *et al.* [18] employed the PIV-LIF method to obtain the velocity field inside and outside the vapour cavity around two hydrofoils. On the basis of the captured images of vapor structures, it can be seen that the cavitation behaves dynamically at the front wall. The backflow causes the separation of cavitation cloud and the vortex remains present inside the separated cloud. The development of unsteady cavitation on impeller blades resembles that on a single hydrofoil. However, in the case of a centrifugal pump, the interaction between the impeller and the volute tongue strongly affects cavitation evolution due to the periodically variation of the flow fields in passages. Limbach *et al.* [19, 20] conducted the numerical calculation and experimental work on a low-specific-speed centrifugal pump. It is observed that the Net positive suction head (*NPSH*) rises toward overload due to incidence, flow separation, and vapor zones at the volute tongue. The numerical simulation has been performed by Micha *et al.* [21] to study the effect of tip-vortex cavitation in a centrifugal pump and its result on the change in nuclei size. The size of bubble nuclei significantly affect the incipient of cavitation and unsteady pressure distribution in passages and the suction surface of the impeller leads to separation and re-circulation at off-design condition which in turn influences the onset of cavitation at the leading edge of the blade. Ahmed *et al.* [22] analyzed the vibration signal in both time and frequency domains to detect and

diagnose the cavitation phenomenon within a centrifugal pump. With various flow rates and rotational speed, the signals presented similar statistical features and revealed that using low-frequency was sensitive to predict cavitation in the pump. By means of unsteady numerical computation, Tang *et al.* [23] found that the blade passing frequency is the dominant frequency of the pressure fluctuations in the casing except the vicinity of the volute tongue for all operating conditions, and the dominant one near the volute tongue is the blade passing frequency at the design point and 0~0.5 times the blade passing frequency at other off-design points. Dönmez *et al.* [24] and Tao *et al.* [25] investigated the influence of geometric parameters of blade inlet on cavitation phenomenon of the centrifugal pump. Cavitation performance of pump is excessively affected by both blade inlet angle and blade leading-edge shape. Increasing hub blade angle has slightly negative effect on cavitation performance of the pump, and the round and ellipse leading-edge impellers have higher inception cavitation coefficient than the ones with blunt or sharp leading-edge. Alex *et al.* [26] compiled an overview on the effect of cavitation in the performance of centrifugal pump. They took account of parameters such as blade numbers, blade angle, inlet flow angle, flow rates, and inlet and outlet pressure. These parameters all have influence on the cavitation performance, however the casing type or the tongue shape were not considered

Above research results indicate that the evolution of cavitation causes the performance deterioration of centrifugal pump and the development of cavitation cloud in passages usually emerges at the suction side of blade leading edge. During the experiment, the pump was operating at overload flow rate and cavitation condition leading to 3 % head drop. Detailed observations of cavity evolution near the tongue of casing were made when no obvious cavitation occurs within the impeller. On the basis of numerical simulation and hydraulic testing, this study explored cavitation structure and flow field in the vicinity of the tongue of the volute casing to provide guidance for the optimization design of centrifugal pumps.

3. Materials and methods

3.1 Design of test stand

The closed testbed is shown in Fig. 1. The pressure fluctuation and cavitation at tongue downstream experiments were conducted on the testbed, except for measuring the external characteristics of the pump model under different flow rates. In Fig.1, impeller and volute were prepared with PMMA. The testbed was equipped with a pressure transmitter, a turbine flowmeter, and a high-frequency pressure sensor for the effective measurement of testing parameters: turn frequency (n), spindle torque (T), mass flow (Q), pressure fluctuation, and inlet and outlet pressures of the pump model (p_{in} and p_{out}). Design parameters of the model pump are shown in Table 1.

Table 1 Design parameters

$Q_d(\text{m}^3/\text{h})$	10
$H(\text{m})$	11
$n(\text{r}/\text{min})$	1450
$P(\text{kW})$	3
Impeller diameter $D_2(\text{mm})$	192
Impeller inlet diameter $D_1(\text{mm})$	54
Impeller hub diameter $d_h(\text{mm})$	20
Impeller width $b_2(\text{mm})$	5
Blade outlet angle β_2	27.5

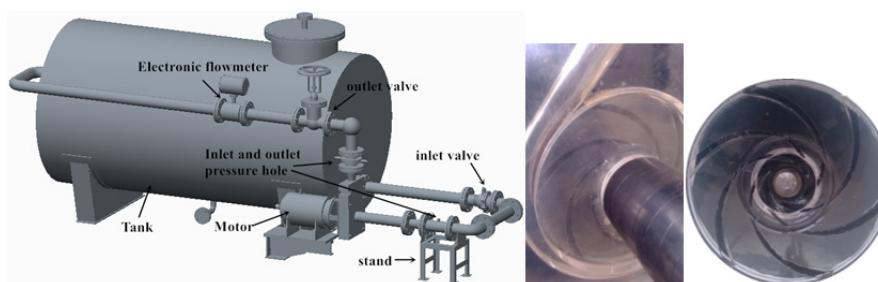


Fig. 1 Three-dimensional schematic of the test stand

3.2 Pressure fluctuation test

The test field of pressure fluctuation is shown in Fig. 2. The overall structure of the testbed is shown in Fig. 2(a). Under the assistance of a high-frequency dynamic pressure sensor, the pressure fluctuation characteristics at the tongue downstream were monitored with a HSJ2010 hydraulic machinery comprehensive tester developed by Huazhong University of Science and Technology. The distributions of pressure fluctuation measurement points are shown in Figs. 2(b) and 2(c). The sampling frequency and sampling time of test data were set to 8,700 Hz and 60 s, respectively.



Fig. 2 The pressure fluctuation testing station

3.3 High-speed photography experiments

The experimental apparatus for high-speed photography is shown in Fig. 3. The positions in Fig. 3 show the locations of the light source and CCD camera in the test. The capture frequency was set to 500 images/s, and the pixel depth of images was set to 8 bits due to the low revolution speed of the test pump. The pixels of the obtained image were 860×1,280. The arrow in Fig. 3 shows the shooting direction of the camera and the incidence angle of light source to decrease the interferences of the reflected light.

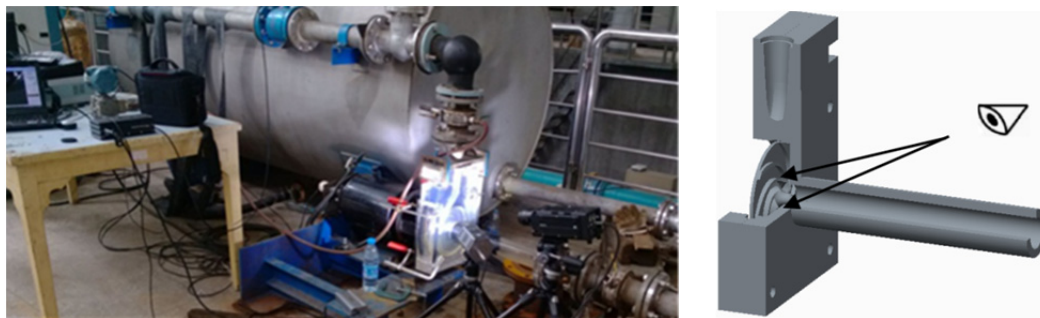


Fig. 3 High-speed photographic experiment device

3.4 Analysis of the test results

The overall external characteristics of the pump are shown in Fig. 4. The flow head curve under non-cavitation is shown in Fig. 4(a), and the *NPSH-H* curve is shown in Fig. 4(b) when $Q/Q_d = 1.52$. With the reduction in *NPSH*, no evident cavitation is found at the back of the impeller inlet. However, evident cloudy cavitation is detected close to the tongue near the flow measuring points.

$$NPSH = \frac{p_{in} - p_v}{\rho g} + \frac{v_{in}^2}{2g} \quad (1)$$

where p_{in} is the static pressure at the impeller inlet, p_v is the saturated vapor pressure, taken as 3574 Pa, v_{in} is the absolute velocity at the impeller inlet.

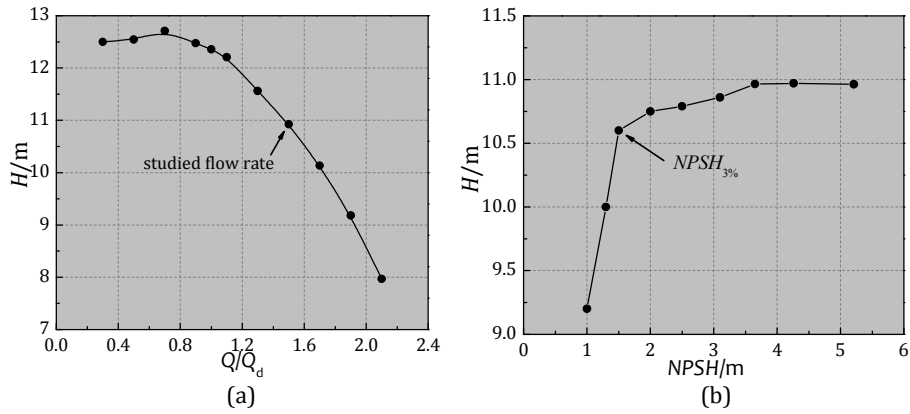


Fig. 4 External characteristic curve of the model pump

The pressure fluctuation in rotating hydraulic machinery can be divided into three types, namely, random pressure, blade multifrequency, and axial multifrequency fluctuations. The random pressure fluctuation is induced by cavitation, eddies, and unsteady secondary flows. It is similar with white noise on the spectrum. The blade frequency fluctuation has a multiple relation with blade passing frequency and is related to rotor-stator interaction. The shaft frequency fluctuation is related to mechanical speed. In this study, the rotational speed of the centrifugal pump was 1,450 rpm. The patterns of time and frequency domains at $Q/Q_d=1.52$ and $NPSH_{3\%}=1.5$ m are shown in Fig. 5. The domain pressure fluctuation in the flow field at the volute downstream is mainly blade passing frequency.

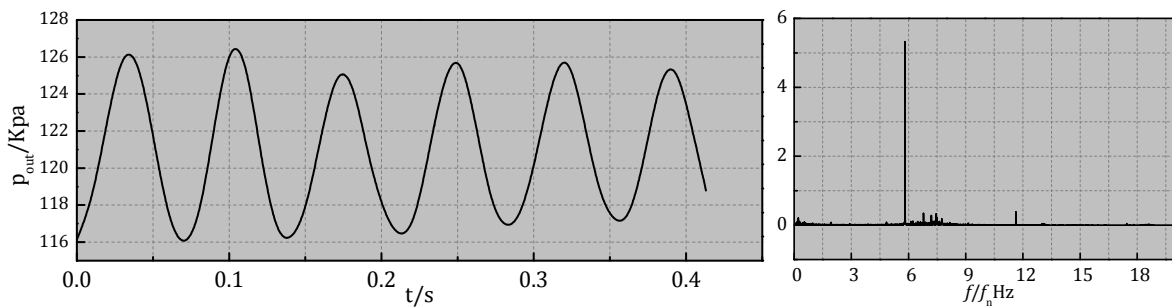


Fig. 5 Time-domain and frequency-domain of the pressure fluctuation in monitoring point

The high-speed photography results at the tongue of the pump model at $Q/Q_d=1.52$ are shown in Fig. 6. Figure 6 shows that: a) A large cavitation structure is developed at the leading edge of the tongue when the blade leaves the observation range, b) The cavitation cloud at the leading edge of the tongue develops quickly and sheds when the next blade appears in the visual frame, c) The flow regime deteriorates, and many bubbles begin to develop at the leading edge when the blade approaches the tongue, d) The cavitation cloud at the tongue tends to attach on the cavities when the blade is close to the tongue position, e) The cavitation cloud rapidly blocks the flow passage near the tongue because the blade leaves the tongue and moves to the downstream areas, significantly influencing the energy transmission in the volute.

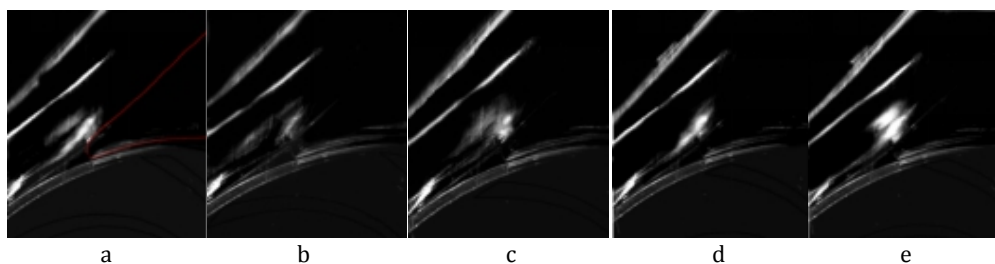


Fig. 6 Cavitation evolution at the volute tongue

4. Result analysis and discussion

4.1 Method of numerical simulation

Time-averaged N-S equations were utilized as the basic governing equations in the single-phase computation. The SST k - ω turbulence model was selected for 3D turbulence numerical simulation because it considers the transmission of turbulence shear stress and can accurately predict the initial position and results of fluid separation under turbulent negative-pressure gradient. The transport behavior was obtained using the eddy viscosity equation containing the limiting quantity

$$v_t = \frac{\alpha_1 k}{\max(a_1 \omega S F_2)} \quad (2)$$

where v_t denotes dynamic viscosity, α_1 is a constant taken as 5/9, k is turbulent kinetic energy, ω is turbulence frequency, F_2 is mixing function that constrains the limiting quantity in the boundary layer, and S is the invariant measure of shear rate.

A homogeneous model was used in the vapor-liquid two-phase flow field. The Zwart equation based on the Rayleigh-Plesset formula was used to analyze the generation and collapse of cavitation bubbles and the mass transfer in the fluid. The evolution process of cavitation bubbles is given as

$$R_B \frac{d^2 R_B}{dt^2} + \frac{3}{2} \left(\frac{dR_B}{dt} \right)^2 + \frac{2\sigma}{\rho_f R_B} = \frac{p_v - p}{\rho_f} \quad (3)$$

where R_B is the radius of bubble, p_v is the pressure inside the bubble, p is the pressure of the fluid around bubble, ρ_f is the density of the fluid, and σ is the surface tension of the interface between the fluid and bubble.

ANSYS ICEM was used to generate high-quality hexahedral meshes in the computational domains of the model pump. Fifteen mesh layers were added to each boundary while guaranteeing that the distribution of blocks was accordance with the flow regime in the computational domains to ensure the accuracy of the numerical simulation in the near-wall zone. Mesh independence was applied under the design condition to guarantee calculation accuracy and improve calculation efficiency. When the grid number exceeded 2.6 million, the change in the pump head was within 1 %. Thus, the grid number was determined. Fig. 7 shows a 3D model of the computational domain, and Fig. 8 shows a schematic of the computation meshes. The left side presents the cross-section mesh of the impeller and volute, and the right side shows partial enlargement of the mesh at the inlet and outlet of blade.

Steady-stage computation of multiple working conditions was performed with different inlet attack angles by using ANSYS-CFX 16.0 software. Given that $NPSH$ is closely related to the pressure at the pump inlet, the total pressure in the inlet and the mass flow in the outlet were used for the computational domain. The calculations for non-cavitation simulation were used as the initial results for the cavitation simulation to reduce the calculation time.

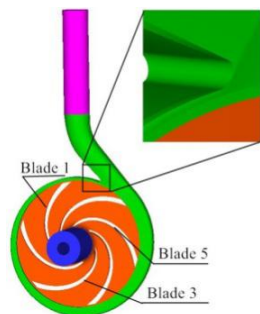


Fig. 7 Computational domain

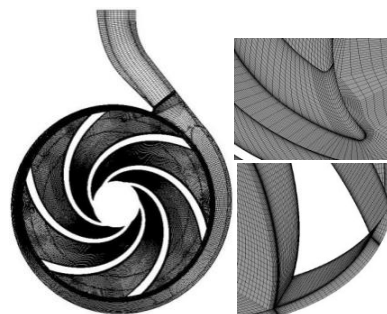


Fig 8 Schematic of the mesh

4.2 Numerical simulation results and analysis

The one-sixth cycle numerical simulation results of cavitation at the tongue at $Q/Q_d = 1.52$ are shown in Fig. 9. The short tongue adopted in this study differs significantly from ordinary airfoil profile but its cavitation structure has great similarities to the attachment and shedding of cavities on single airfoil. The inlet speed and Strouhal number are the main factors influencing periodic cavitation formation and shedding on single airfoil. For the cavitation at the tongue, the development of cavitation is influenced by the interaction changes in pressure gradients between the pressure surface (PS) and suction surface (SS) of blades and the unsteady changes in flow field near the leading edge of the tongue caused by jet wake at the impeller outlet.

As shown in Fig. 9(a), the tongue at 0 T is in the middle of flow passage. At this moment, a relatively large cavitation cloud occurs at the leading edge of the tongue. At $1/4 T$, the cavitation cloud develops quickly at the tongue and sheds when the next blade approaches the tongue. At $1/2 T$, the cavitation cloud gradually moves toward the tongue downstream when the blade approaches the tongue. The cavitation cloud separates in the high-pressure zone at the tongue downstream, thereby decreasing the area of low-pressure zone. At $3/4 T$, the blade is located at the tongue, and a newly attached cavitation emerges at the tongue. The previous cavitation cloud breaks quickly and then moves downward. The flow regime in the volute deteriorates, and cavitation wake occurs at the downstream close to the volute outlet. Thus, the low-pressure zone is expanded. At $1 T$, the blade leaves the tongue, and the cavitation cloud develops at the leading edge of the tongue, exerting the evident blocking effect of the flow passage. With the shedding of cavities, the actual flow area close to the tongue increases suddenly, thereby causing great energy losses in the volute.

The velocity diagram for the middle section of the pump in a single cycle is shown in Fig. 9(b). Uniform relative rates are found on the middle stream surface of the blade, and a local high-pressure zone is found at the blade outlet close to the tongue. At overload conditions, serious separation occurs at the tongue. The negative pressure fluctuation caused by separation and transition can trigger the initiation of cavitation. Separated unsteady eddies are formed near the tongue, which are the vibration and noise sources. At 0 T, a low-rate backflow zone that occupies approximately one-half of the flow passage is formed at the leading edge and tongue downstream, significantly blocking the flow passage. The cavitation cloud develops quickly and squeezes the actual flow area. At $1/4 T$ and $1/2 T$, the blocked passage is released partially with the shedding of cavities. The low-rate backflow zone shrinks, and the eddies move downward. The flow field near the tongue is divided into backflow and mainstream zones. At $3/4 T$, the blade outlet is close to the tongue, and the eddy center moves downward. The previous cavitation cloud moves downward to form a cavitation wake. At $1 T$, the blade leaves the tongue, and a new cavitation cloud emerges at the leading edge of the tongue.

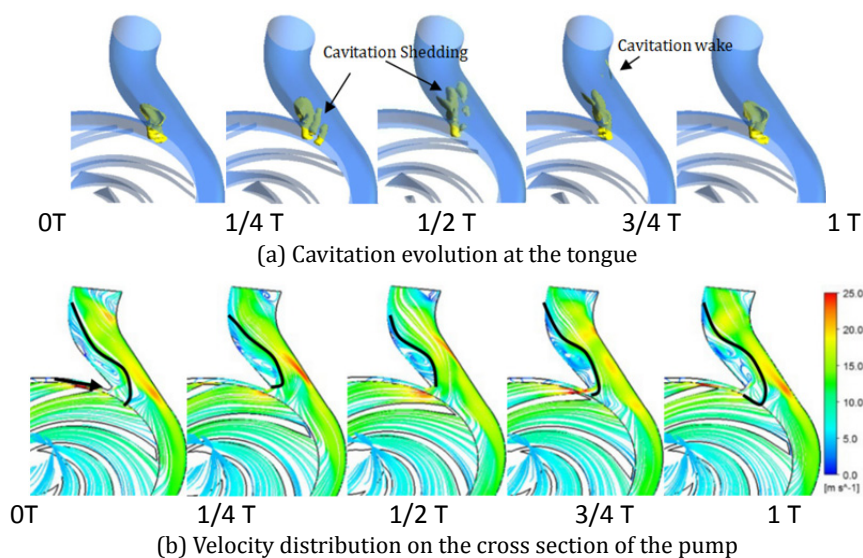


Fig. 9 Flow field at $Q/Q_d = 1.52$

A centrifugal pump with a helical pumping chamber can generate a radial force that acts on the impeller during operation. Thus, the axis bears alternating stresses and generates a directed deflection. The pressure in the chamber of the centrifugal pump continuously changes due to the interference by the impeller and the tongue of the volute, thereby generating unsteady radial forces. The existence of cavitation may affect the radial forces. The vector diagram of radial force distribution on the impeller in a single cycle under different flow rates is shown in Fig. 10. Cavitation occurs at the impeller inlet when $Q/Q_d = 1.3$ and 1.52. However, the radial force distributions at the critical cavitation state are basically consistent under three working conditions. The radial force distribution is related to the number of blades and is in hexagonal star distribution. This condition is caused by the interference of the impeller and the volute. At $Q/Q_d = 1.3$, the vector diagrams under critical cavitation and non-cavitation conditions are close. At $Q/Q_d = 1.52$ and 1.73, the radial forces produce alternating stresses under cavitation conditions when the blade sweeps the separation tongue.

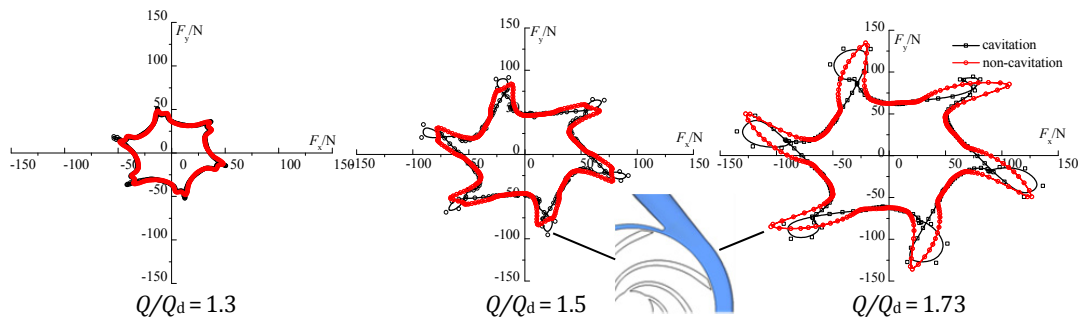


Fig. 10 Distribution of radical force

The secondary flows and “jet wake” phenomenon are caused by the high-speed rotation of the impeller, interference between the impeller and the volute, and the viscosity of fluid due to the spatial asymmetric structure of the centrifugal pump. The internal flow field presents complicated unsteady characteristics that may cause pressure fluctuation. The circumferential pressure fluctuation of the volute and pressure fluctuation near the tongue are influenced and present different characteristics when cavitation emerges at the tongue. The monitoring points of pressure fluctuation in the volute are shown in Fig. 11.

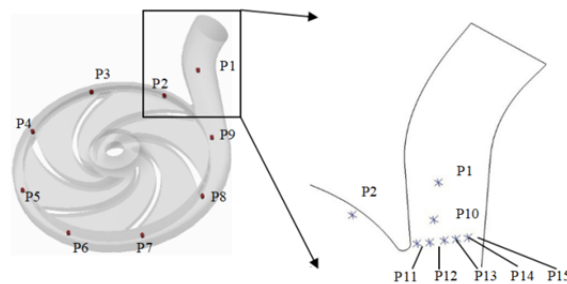


Fig. 11 Monitoring points of pressure fluctuation

The pressure amplitude of pressure pulsation in the volute is characterized, and the influences of the static pressure of monitoring points on pressure fluctuation are eliminated. The strength of pressure fluctuation is expressed as pressure coefficient C_p

$$C_p = \frac{p - \bar{p}}{\frac{1}{2}\rho U^2} \quad (4)$$

where p is the instantaneous pressure value at the monitoring points, and \bar{p} is the mean pressure of monitoring points in the investigated cycle.

The time domain of monitoring points P1-P9 when $Q/Q_d = 1.3, 1.52,$ and 1.73 is shown in Fig. 12. The pressure fluctuation intensity increases significantly with the increase of flow rate. Although the number of wave peaks at different monitoring points in a single cycle is equal to the number of blades, the regularity in a single cycle weakens with the increase of flow rate. At $Q/Q_d =$

1.3, the pressure fluctuation at different monitoring points shows evident periodic fluctuation laws. The fluctuation amplitudes at different monitoring points are close, except for the monitoring points at the downstream of the tongue. Considering that all monitoring points are close to the flow field at the impeller outlet, the pressure fluctuation is caused by the jet-wake structure at the passage outlet of the impeller, and the fluctuation amplitudes are similar. At $Q/Q_d=1.73$, irregular wave peaks are found, indicating the occurrence of serious cavitation close to the tongue. Therefore, the flow field becomes extremely disordered.

The multiplication of rotating frequency is defined as

$$NF = \frac{60F}{n} = \frac{F}{F_n} \quad (5)$$

where F is the practical frequency after Fourier transform, n is the rotating speed of the impeller, and F_n is the rotating frequency under the corresponding rotating speed.

As shown in Fig. 13, the excitation frequency at different monitoring points is the blade passing frequency under different flow rates. The shaft frequency and other low-frequency and high-frequency bands have small amplitudes. This finding reveals that the pressure fluctuation caused by the “jet wake” flow structure under high flow rate is the main excitation frequency. With the increase in flow rate, the pressure fluctuation strength increases significantly. In particular, a magnitude of jumps of pulsation strength is found after the cavitation occurs at the tongue. At $Q/Q_d=1.3$, the pressure fluctuation is weak. The main excitation frequency first decreases and then increases from the first to the eighth sections. The pressure fluctuation strength declines as the monitoring point moves away from the tongue. At $Q/Q_d=1.52$ and 1.73 , the cavitation in the volute becomes evident, and the pressure fluctuation in the volute shows different laws. The pressure fluctuation at the downstream of the tongue is lower than that at the impeller outlet. The fluctuation strength at the impeller outlet decreases gradually from the first to the eighth sections.

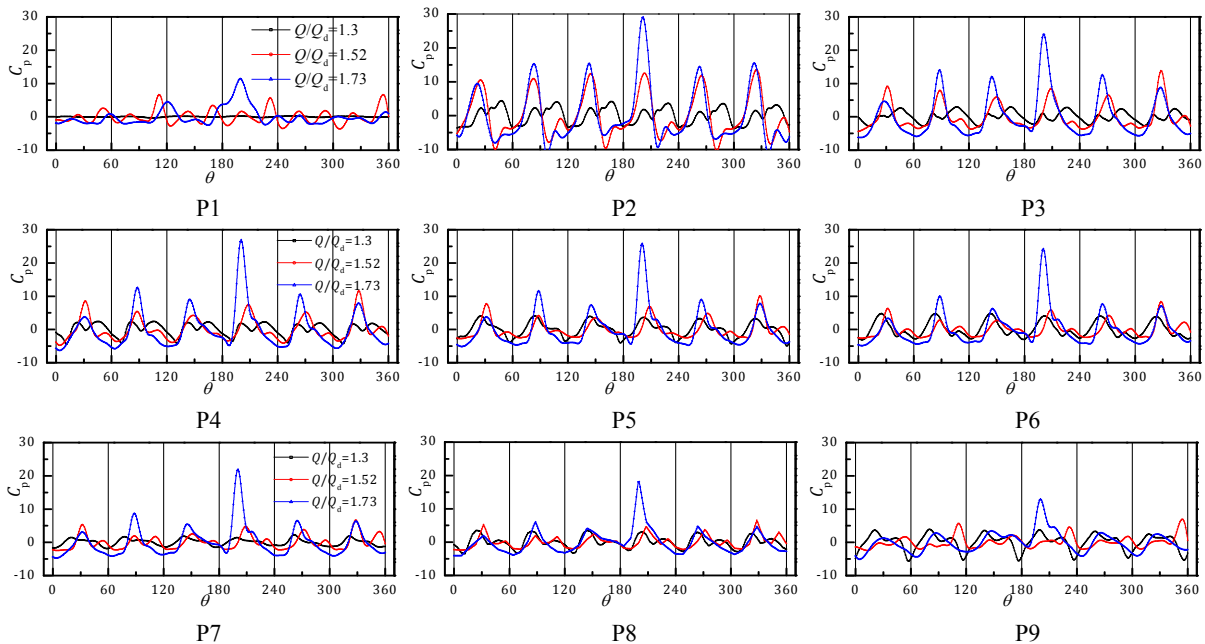


Fig. 12 Pressure fluctuation of monitoring points around the impeller outlet

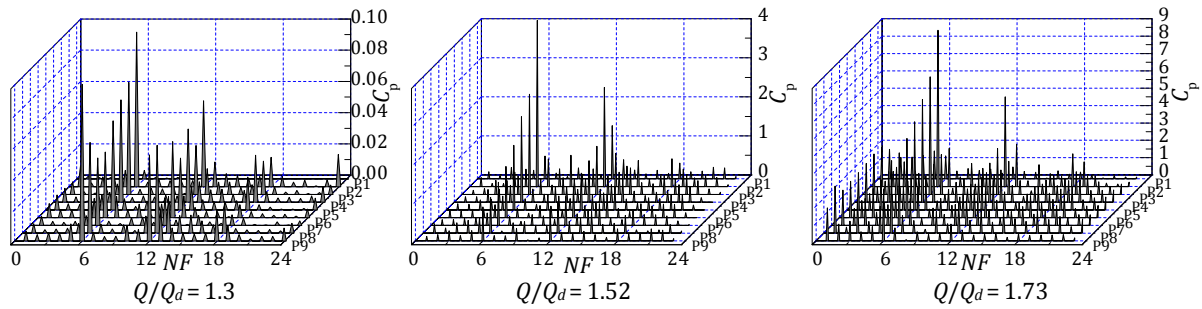


Fig. 13 Frequency domain of pressure fluctuation

The time domains of pressure fluctuation at P10-P15 are shown in Fig. 14. The pressure fluctuation has six wave peaks under different working conditions, and the pressure fluctuation at different monitoring points is caused by the interference of the tongue. The pressure fluctuation strength increases with the increase of flow rate. The fluctuation strength weakens when the distance between the monitoring point and the tongue increases. The influence of the tongue on pressure fluctuation and the influence of jet-wake structure at the impeller outlet decline when the monitoring point approaches the wall surface facing the tongue. The pressure fluctuation strength without cavitation close to the tongue at $Q/Q_d = 1.3$ is significantly lower than that with cavitation at the tongue at $Q/Q_d = 1.52$ and 1.73 . This finding reflects that the unsteady flow at the tongue is complicated, and the changes in pressure distribution are intense after the occurrence of cavitation. At actual operation, these conditions may trigger serious vibration noises, thereby influencing the stable operation of the pump.

The frequency domains of pressure fluctuations at different monitoring points close to the tongue under different working conditions are shown in Fig. 15. The dominant frequency under working conditions is the blade passing frequency. The rapid increase in dominant frequency amplitude caused by cavitation reflects the influences of cavitation development on the pressure fluctuation strength close to the tongue. The pressure fluctuation at the downstream monitoring point P1 is 0.41 of the dominant frequency amplitude under non-cavitation condition. However, it increases to 2.12 and 3.84 when $Q/Q_d = 1.52$ and $Q/Q_d = 1.73$, respectively. A strong pressure fluctuation may be found at the downstream position of the tongue due to the development and breakage of cavities. In accordance with the changes in dominant frequency amplitude under different working conditions, the pressure fluctuation intensity declines gradually when the monitoring point approaches the wall surface facing the separation tongue.

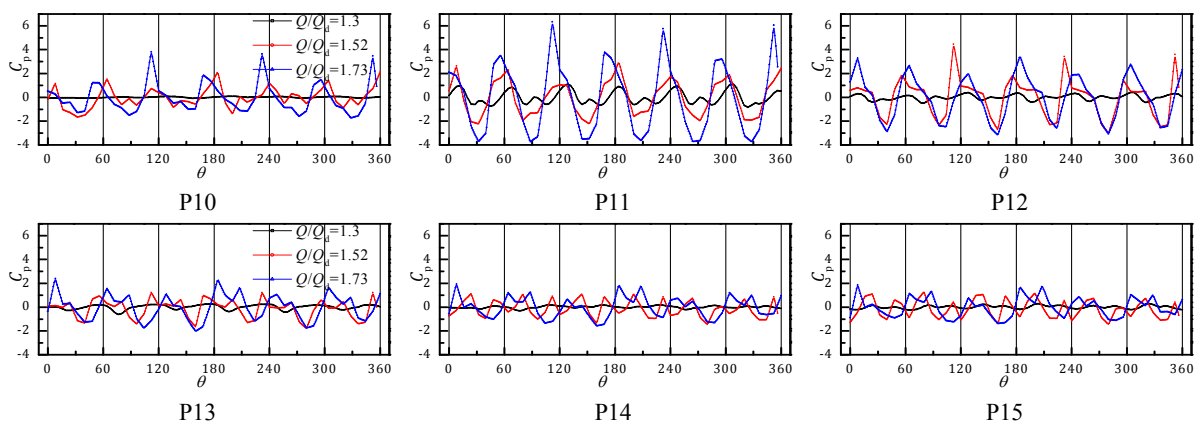


Fig. 14 Pressure fluctuation near the tongue

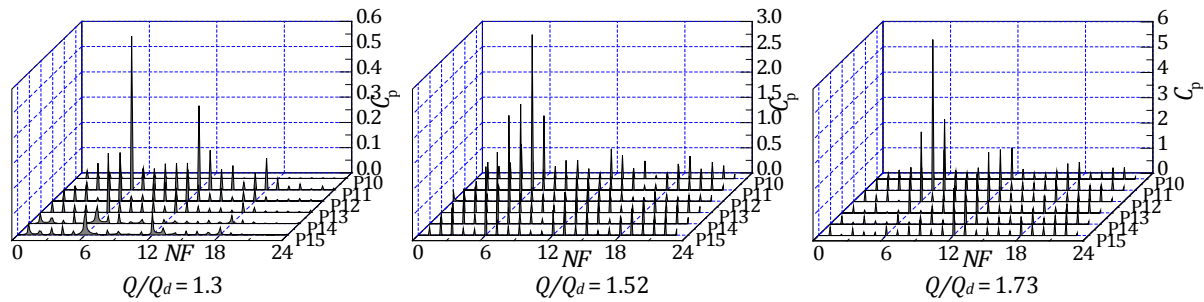


Fig. 15 Frequency domain of pressure fluctuation of monitoring points near the tongue

5. Conclusion

In this study, an attempt is made to investigate the cavitation at the tongue of centrifugal pump at overload conditions. By means of numerical computation and visualization measurement, unsteady flow structures in cavitation zone were studied to associate with blade loading and pressure fluctuation. From the present research following conclusions can be drawn:

- As the increase of flow rate, the shedding of the separated vortex lowers the static pressure near the tongue. And the flow separation at the tongue brings dramatic pressure fluctuation and strong shearing vortex which induce cavitation.
- At overload conditions, higher radical velocity and deviation of relative flow angle contribute to periodical variation of flow field near the tongue, which results in periodical cavitation evolution. Thus, the frequency of the cavitation cloud shedding is equal to the blade passing frequency.
- Cavitation at the tongue not only enhances the pressure fluctuation in volute, but also affects the blade loading distribution: The pressure pulsation in the volute is consistent with the blade passing frequency whether cavitation occurs or not, while the pulsation intensity increases obviously after cavitation inception. From the first section to the eighth section of volute, the pulsation intensity of impeller outlet decreases gradually.

The study showed that it is possible that the head drop at overload conditions is caused by the appearance of cavitation at the tongue of volute, which may provide guidance for the optimization of the anti-cavitation performance of centrifugal pump. In future, more accurate visualization measurement and vibration experiment can be conducted to investigate the cavitation phenomenon at the tongue of casing.

Acknowledgement

This research was funded by the National Key R&D Program of China (2018YFC0810506) and the Key R&D Program of Zhenjiang (SH2017049).

References

- [1] Brennen, C.E. (2011). *Hydrodynamics of Pumps*, Cambridge University Press, Cambridge, United Kingdom, doi: [10.1017/CBO9780511976728](https://doi.org/10.1017/CBO9780511976728).
- [2] Franc, J.-P., Michel, J.-M. (2005). *Fundamentals of cavitation*, Springer, New York, USA, doi: [10.1007/1-4020-2233-6](https://doi.org/10.1007/1-4020-2233-6).
- [3] Širok, B., Dular, M., Hočevar, M., Novak, M., Stoffel, B., Ludwig, G., Bachert, B. (2002). The influence of cavitation structures on the erosion of a symmetrical hydrofoil in a cavitation tunnel, *Strojniški Vestnik – Journal of Mechanical Engineering*, Vol. 48, No. 7, 368-378.
- [4] Wang, S.-L., Tan, L., Wang, Y.-C. (2013). Characteristics of transient cavitation flow and pressure fluctuation for a centrifugal pump, *Journal of Vibration and Shock*, Vol. 32, No. 22, 168-173, doi: [10.13465/j.cnki.jvs.2013.22.024](https://doi.org/10.13465/j.cnki.jvs.2013.22.024).
- [5] Bilus, I., Predin, A. (2009). Numerical and experimental approach to cavitation surge obstruction in water pump, *International Journal of Numerical Methods for Heat & Fluid Flow*, Vol. 19, No. 7, 818-834, doi: [10.1108/09615530910984091](https://doi.org/10.1108/09615530910984091).
- [6] Lu, J., Yuan, S., Ren, X., Liu, Y., Si, Q. (2015). Investigation of instabilities of cavitation at low flow rate of centrifugal pump, *Transactions of the Chinese Society for Agricultural Machinery*, Vol. 46, No. 8, 54-58, doi: [10.6041/j.issn.1000-1298.2015.08.009](https://doi.org/10.6041/j.issn.1000-1298.2015.08.009).

- [7] Čudina, M., Prezelj, J., Černetič, J. (2012). Use of noise and vibration spectra to detection cavitation in kinetic pumps, In: *Proceedings of the 8th International Symposium on Cavitation*, Singapore, 978-971, doi: [10.3850/978-981-07-2826-7_172](https://doi.org/10.3850/978-981-07-2826-7_172).
- [8] Lu, Z., He, X., Wang, C. (2018). Influencing factors of self-priming time of multistage self-priming centrifugal pump, *DYNA – Ingeniería e Industria*, Vol. 93, No. 6, 630-635, doi: [10.6036/8930](https://doi.org/10.6036/8930).
- [9] Ključanin, D., Manduka, A. (2019). The cantilever beams analysis by the means of the first-order shear deformation and the Euler-Bernoulli theory, *Tehnički Glasnik*, Vol. 13, No. 1, 63-67, doi: [10.31803/tg-20180802210608](https://doi.org/10.31803/tg-20180802210608).
- [10] Luo, X., Zhang, Y., Peng, J., Xu, H. (2008). Effect of impeller inlet geometry on centrifugal pump cavitation performance, *Journal of Tsinghua University (Science and Technology)*, Vol. 48, No. 5, 836-839, doi: [10.3321/j.issn:1000-0054.2008.05.019](https://doi.org/10.3321/j.issn:1000-0054.2008.05.019).
- [11] Castorani, V., Landi, D., Mandolini, M., Germani, M. (2019). CFD simulations of filter houses for power plant gas turbine: Evaluation of differences between 2D and 3D models, *DYNA – Ingeniería e Industria*, Vol. 94, 145-149, doi: [10.6036/8901](https://doi.org/10.6036/8901).
- [12] Blecich, P., Senčić, T., Wolf, I., Bonefačić, I. (2018). Numerical investigation of heat and mass transfer inside a wet cooling tower, *Tehnički Glasnik*, Vol. 12, No. 3, 131-138, doi: [10.31803/tg-20171017145907](https://doi.org/10.31803/tg-20171017145907).
- [13] Kunz, R.F., Boger, D.A., Stinebring, D.R., Chyczewski, T.S., Lindau, J.W., Gibeling, H.J., Venkateswaran, S., Govindan, T.R. (2000). A preconditioned Navier–Stokes method for two-phase flows with application to cavitation prediction, *Computers & Fluids*, Vol. 29, No. 8, 849-875, doi: [10.1016/s0045-7930\(99\)00039-0](https://doi.org/10.1016/s0045-7930(99)00039-0).
- [14] Luo, X., Wei, W., Ji, B., Pan, Z., Zhou, W., Xu, H. (2013). Comparison of cavitation prediction for a centrifugal pump with or without volute casing, *Journal of Mechanical Science and Technology*, Vol. 27, No. 6, 1643-1648, doi: [10.1007/s12206-013-0411-5](https://doi.org/10.1007/s12206-013-0411-5).
- [15] Bachert, R., Stoffel, B., Dular, M. (2010). Unsteady cavitation at the tongue of the volute of a centrifugal pump, *Journal of Fluids Engineering*, Vol. 132, No. 6, Article No. 061301, doi: [10.1115/1.4001570](https://doi.org/10.1115/1.4001570).
- [16] Xue, M., Park, Y. (2012). Large eddy simulations of separated flow at the casing tongue of an afterburner fuel pump, *Journal of Tsinghua University (Science and Technology)*, Vol. 52, No. 11, 1638-1642, doi: [10.16511/j.cnki.qhdxxb.2012.11.023](https://doi.org/10.16511/j.cnki.qhdxxb.2012.11.023).
- [17] Meng, L., He, M., Zhou, L., Yang, J., Wang, Z., Karney, B. (2016). Influence of impeller-tongue interaction on the unsteady cavitation behavior in a centrifugal pump, *Engineering Computations*, Vol. 33, No. 1, 171-183, doi: [10.1108/EC-09-2014-0179](https://doi.org/10.1108/EC-09-2014-0179).
- [18] Dular, M., Širok, B., Bachert, R., Stoffel, B. (2005). Transient simulation, visualization and PIV-LIF measurements of the cavitation on different hydrofoil configurations, *Strojniški Vestnik – Journal of Mechanical Engineering*, Vol. 51, No. 1, 13-27.
- [19] Limbach, P., Kimoto, M., Deimel, C., Skoda, R. (2014). Numerical 3D simulation of the cavitating flow in a centrifugal pump with low specific speed and evaluation of the suction head, In: *Proceedings of the ASME Turbo Expo 2014: Turbine Technical Conference and Exposition. Volume 2D: Turbomachinery*, Düsseldorf, Germany, 16-20, doi: [10.1115/GT2014-26089](https://doi.org/10.1115/GT2014-26089).
- [20] Limbach, P., Skoda, R. (2017). Numerical and experimental analysis of cavitating flow in a low specific speed centrifugal pump with different surface roughness, *Journal of Fluids Engineering*, Vol. 139, No. 10, Article No. 101201, doi: [10.1115/1.4036673](https://doi.org/10.1115/1.4036673).
- [21] Micha Premkumar, T., Sathish Babu, R., Vinoth Kumar, M., Hariram, V., Seralathan, S. (2019). Numerical simulation of hydrodynamic cavitation in centrifugal pump, *International Journal of Innovative Technology and Exploring Engineering*, Vol. 8, No. 11, 2689-2693, doi: [10.35940/ijitee.k2140.0981119](https://doi.org/10.35940/ijitee.k2140.0981119).
- [22] Al-Obaidi, A.R. (2019). Investigation of effect of pump rotational speed on performance and detection of cavitation within a centrifugal pump using vibration analysis, *Heliyon*, Vol. 5, No. 6, Article No. E01910, doi: [10.1016/j.heliyon.2019.e01910](https://doi.org/10.1016/j.heliyon.2019.e01910).
- [23] Tang, X., Zou, M., Wang, F., Li, X., Shi, X. (2017). Comprehensive numerical investigations of unsteady internal flows and cavitation characteristics in double-suction centrifugal pump, *Mathematical Problems in Engineering*, Vol. 2017, Article ID 5013826, doi: [10.1155/2017/5013826](https://doi.org/10.1155/2017/5013826).
- [24] Dönmez, A.H., Yumurtacı, Z., Kavurmacioğlu, L. (2018). The effect of inlet blade angle variation on cavitation performance of a centrifugal pump: A parametric study, *Journal of Fluids Engineering*, Vol. 141, No. 2, Article No. 021101, doi: [10.1115/1.4040557](https://doi.org/10.1115/1.4040557).
- [25] Tao, R., Xiao, R., Wang, Z. (2018). Influence of blade leading-edge shape on cavitation in a centrifugal pump impeller, *Energies*, Vol. 11, No. 10, Article No. 2588, doi: [10.3390/en11102588](https://doi.org/10.3390/en11102588).
- [26] George, A., Muthu, P. (2015). Review on effects of cavitation in the performance of centrifugal pump, *International Journal of Advance Engineering and Research Development*, Vol. 2, No. 11, 298-303, doi: [10.21090/ijaerd.021148](https://doi.org/10.21090/ijaerd.021148).

Optimization of process performance by multiple pentagon fuzzy responses: Case studies of wire-electrical discharge machining and sputtering process

Al-Refaie, A.^a, Lepkova, N.^{b,*}, Abbasi, G.^a, Bani Domi, G.^a

^aUniversity of Jordan, Department of Industrial Engineering, Jordan

^bVilnius Gediminas Technical University, Department of Construction Management and Real Estate, Lithuania

ABSTRACT

This research developed mathematical models to optimize process performance for multiple pentagon fuzzy quality responses. Initially, each quality response was represented by a pentagon membership function. Then, the combination of optimal factor levels was obtained for each response replicate. Those optimal combinations were then used to construct pentagon regression models for each response. A pentagon fuzzy optimization model was formulated and solved to determine the combination of optimal factor levels at each element of pentagon response's fuzzy number. Two real case studies, i.e. wire-electrical discharge machining and sputtering process, were provided for illustration. Optimal results of the two case studies revealed that the proposed procedure effectively optimized performance under uncertainty and provided larger improvement in multiple quality characteristics. In conclusion, the proposed procedure may enhance the process engineer's knowledge about effects of uncertainty on process/product performance and help practitioners decide the proper adjustments of factor levels in order to enhance performance of electrical discharge machining and sputtering process.

© 2020 CPE, University of Maribor. All rights reserved.

ARTICLE INFO

Keywords:

Modeling and optimization;
Fuzzy goal programming;
Pentagon regression modelling;
Pentagon fuzzy numbers;
Wire electro-discharge machining (WEDM);
Surface roughness (SR);
Material removal rate (MRR);
Sputtering process;
Gallium-doped ZnO (GZO)

*Corresponding author:

natalija.lepkova@vgtu.lt
(Lepkova, N.)

Article history:

Received 7 August 2020
Revised 16 October 2020
Accepted 19 October 2020

1. Introduction

In practice, inherent variations in a manufacturing process and measurement system are unavoidable. When conducting designed experiments, such variations may result in obtaining erroneous combination of optimal process factor settings, and thereby may not lead to produce product/process improvement as expected [1]. Most literature studies ignored the effect on uncertainty and provided a certain combination of optimal factor settings [2, 3]. Hocine *et al.* [4] extended the conventional fuzzy goal programming model to solve a wide range of uncertainties decision-making problems. The model was validated by optimizing the renewable portfolio for electricity generation in Italy. Komsiyah *et al.* [5] analysed production planning problem in a furniture Company with different operational constraint, including production time, quantity of raw materials, and warehouse capacity. The fuzzy goal programming was applied to minimize the production cost and raw material cost, and maximize the profit. Mirzaee *et al.* [6] emphasized the problem of supplier selection, which was mathematically formulated by a mixed integer linear programming model. This model was then solved by a pre-emptive fuzzy goal pro-

gramming approach. Johnson and Bogle [7] presented a model-based approach to risk analysis to aid design for pharmaceutical processes which combined systematic modelling procedures with Hammersley sampling-based uncertainty analysis and sensitivity analysis used to quantify predicted performance uncertainty and to identify key uncertainty contributions. Authors demonstrated the methodology on an industrial case study where the process flowsheet was fixed and some pilot data was available.

In order to enhance process knowledge about the influence of such variations on process/product performance, an effective optimization procedure is required. The fuzzy goal programming (FGP) is found an effective technique in handling the challenge of optimizing process performance under responses' uncertainty [8-11]. Further, Al-Refaie *et al.* [12] expressed each quality characteristic by triangular fuzzy number and then used fuzzy regression combined with desirability function to optimize process performance. However, in order to obtain better evaluation of process performance under fuzziness over a wider range of optimal factor levels and guide process engineers on taking proper adjustment of key factor levels, this research treats each quality response as a pentagon fuzzy number. Then, a pentagon fuzzy regression-desirability procedure is developed to optimize process performance. The remaining of this paper is organized as follows. Section two defines materials and methods, Section three presents results and discussion, finally, section four summarizes conclusion.

2. Materials and methods

2.1 Optimization procedure

Assume a production process is studied via designed experiments. Typically, at each combination of process controllable factors product samples are collected and then observations of each quality characteristic's replicate are recorded. Then, the proposed procedure to optimal process performance for fuzzy multiple quality responses goes as follows:

Step I: Formulate the multiple regression model between the $y_{jr}(x)$; the response value of the r -th replicate of response j , and controllable process factors, x , as shown in Eq. 1.

$$y_{jr}(x) = \beta_{0r} + \sum_{f=1}^v \beta_{fr} x_f + \sum_{f=1}^v \beta_{ffr} x_f^2 + \sum_{g < f} \beta_{fgr} x_f x_g + \varepsilon, \quad r = 1, 2, \dots, k \quad (1)$$

where β_{0r} is the intercept and the coefficients, β_f, β_{fg} , and β_{ff} , are crisp values, while, x_f, x_f^2 , and $x_f x_g$ denote independent factor variables and ε is random error.

Utilizing the intercept and coefficients for each response replicate, obtained in Eq. 1, the fuzzy multilinear regression is developed for each fuzzy response, $y_j(\tilde{x})$, as follows:

$$\tilde{y}_j(\tilde{x}) = \tilde{\beta}_0 + \sum_{f=1}^v \tilde{\beta}_f \tilde{x}_f + \sum_{f=1}^v \tilde{\beta}_{ff} \tilde{x}_f^2 + \sum_{g < f} \tilde{\beta}_{fg} \tilde{x}_f \tilde{x}_g + \varepsilon \quad \forall j, \forall f \quad (2)$$

where $\tilde{\beta}_0, \tilde{\beta}_{ff}$, and $\tilde{\beta}_{fg}$ are pentagonal fuzzy coefficients ($\beta^a, \beta^b, \beta^c, \beta^d, \beta^e$) and calculated as:

$$\tilde{\beta} = \begin{cases} \beta^a = \beta^c - s \\ \beta^b = \beta^c - \lambda s \\ \beta^c = \text{Average}(\beta_1, \dots, \beta_k) \\ \beta^d = \beta^c + \lambda s \\ \beta^e = \beta^c + s \end{cases} \quad (3)$$

where s is the estimated standard deviation of β values and λ is a constant between zero and one.

Step II: Represent each of the quality responses and process factors by adequate membership functions from Fig. 1, MSFs, as follows:

- a) For the nominal-the-best (NTB) type quality response, the negative deviations, d_1^- and d_2^- , and positive deviations, d_1^+ and d_2^+ , from the target, T , are the decision variables. The maximal negative admissible violation, D_1^- and D_2^- , and positive admissible violation, D_1^+ and D_2^+ , from T are the parameters.

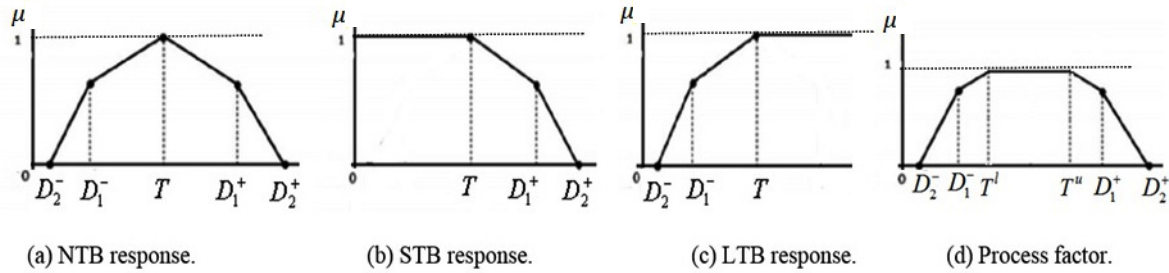


Fig. 1 The selected membership functions for quality responses

The objective function for the NTB type response is to minimize the sum of the weighted positive and negative deviations as shown in Eq. 4a.

$$\text{Min} \sum_{\forall NTB} \frac{d_1^-}{D_1^-} + \frac{d_2^-}{D_2^-} + \frac{d_1^+}{D_1^+} + \frac{d_2^+}{D_2^+} \quad [\text{Minimize sum of deviation ratios}] \quad (4a)$$

The objective function is subject to the following constraints:

$$y + d_1^+ + d_2^+ - d_1^- - d_2^- = T \quad [\text{Response target}] \quad (4b)$$

$$\mu + \frac{d_1^-}{D_1^-} + \frac{d_2^-}{D_2^-} + \frac{d_1^+}{D_1^+} + \frac{d_2^+}{D_2^+} = 1 \quad [\text{Response membership value}] \quad (4c)$$

$$0 \leq d_1^- \leq D_1^-, \quad 0 \leq d_2^- \leq D_2^- \quad [\text{Negative deviation ranges}] \quad (4d)$$

$$0 \leq d_1^+ \leq D_1^+, \quad 0 \leq d_2^+ \leq D_2^+ \quad [\text{Positive deviation ranges}] \quad (4e)$$

b) For the smaller-the-better (STB) type quality response, the objective function is to minimize the sum of weighted positive deviations:

$$\text{Min} \sum_{\forall STB} \frac{d_1^+}{D_1^+} + \frac{d_2^+}{D_2^+} \quad [\text{Minimize sum of positive deviation ratios}] \quad (5a)$$

Subject to:

$$y - d_1^+ - d_2^+ \leq T \quad [\text{Response target}] \quad (5b)$$

$$\mu + \frac{d_1^+}{D_1^+} + \frac{d_2^+}{D_2^+} = 1 \quad [\text{Response membership value}] \quad (5c)$$

$$0 \leq d_1^+ \leq D_1^+, \quad 0 \leq d_2^+ \leq D_2^+ \quad [\text{Positive deviation ranges}] \quad (5d)$$

c) For the larger-the-better (STB) type response, the objective function is to minimize the sum of weighted negative deviations.

$$\text{Min} \sum_{\forall LTB} \frac{d_1^-}{D_1^-} + \frac{d_2^-}{D_2^-} \quad [\text{Minimize sum of positive deviation ratios}] \quad (6a)$$

Subject to:

$$y + d_1^- + d_2^- \geq T \quad [\text{Response target}] \quad (6b)$$

$$\mu + \frac{d_1^-}{D_1^-} + \frac{d_2^-}{D_2^-} = 1 \quad [\text{Response membership value}] \quad (6c)$$

$$0 \leq d_1^- \leq D_1^-, \quad 0 \leq d_2^- \leq D_2^- \quad [\text{Negative deviation ranges}] \quad (6d)$$

d) Each process factor, x , is represented by a trapezoidal MSF with preferable upper and lower limits, T^l and T^u , respectively. Let the maximal negative admissible violations from T^l be denoted as D_1^- and D_2^- while the positive admissible violations from T^u be denoted as D_1^+ and D_2^+ . The objective function is to minimize the sum of the weighted positive and negative deviations and is formulated mathematically as shown in Eq. 7a.

$$\text{Min} \sum_{\forall x} \frac{d_1^-}{D_1^-} + \frac{d_2^-}{D_2^-} + \frac{d_1^+}{D_1^+} + \frac{d_2^+}{D_2^+} \quad [\text{Minimize sum of positive deviation ratios}] \quad (7a)$$

Subject to:

$$x + d_1^- + d_2^+ \geq T^l \quad [\text{Factor target}] \quad (7b)$$

$$x - d_1^+ - d_2^- \leq T^u \quad [\text{Factor target}] \quad (7c)$$

$$\mu + \frac{d_1^-}{D_1^-} + \frac{d_2^-}{D_2^-} + \frac{d_1^+}{D_1^+} + \frac{d_2^+}{D_2^+} = 1 \quad \text{[Factor membership value]} \tag{7d}$$

$$0 \leq d_1^- \leq D_1^-, \quad 0 \leq d_2^- \leq D_2^- \quad \text{[Negative deviation ranges]} \tag{7e}$$

$$0 \leq d_1^+ \leq D_1^+, \quad 0 \leq d_2^+ \leq D_2^+ \quad \text{[Positive deviation ranges]} \tag{7f}$$

Finally, the objective function for the full optimization model minimizes the sum of ratios of positive and negative deviations subject to the set of constraints for all responses and process factors. Solving the complete optimization model for each response’s replicate, the combination of optimal factor settings, x^* , can then be obtained for all response replicates.

Step III: Let $\tilde{y}_j(\tilde{x}^q)$ be the value of response j ; $j = 1, \dots, Q$ obtained by substituting optimal factor levels, \tilde{x}^q , for quality characteristic q ; $q = 1, \dots, Q$. Calculate the $\tilde{y}_j(\tilde{x}^q)$ values for all q responses utilizing Eq. (2). Let $\tilde{d}_j(\tilde{y}_j(\tilde{x}^q))$ denotes the fuzzy desirability function are defined as:

$$\text{NTB response} \tag{8a}$$

$$\tilde{d}_j(\tilde{y}_j(\tilde{x}^q)) = \begin{cases} 0, & \tilde{y}_j(\tilde{x}^q) \leq y_{\min} \\ \frac{\tilde{y}_j(\tilde{x}^q) - y_{\min}}{y_{\max} - y_{\min}}, & y_{\min} \leq \tilde{y}_j(\tilde{x}^q) \leq y_{\max} \\ 1, & \tilde{y}_j(\tilde{x}^q) \geq y_{\max} \end{cases} \quad \text{LTB response} \tag{8b}$$

$$\tilde{d}_j(\tilde{y}_j(\tilde{x}^q)) = \begin{cases} 1, & \tilde{y}_j(\tilde{x}^q) \leq y_{\min} \\ \frac{\tilde{y}_j(\tilde{x}^q) - y_{\max}}{y_{\min} - y_{\max}}, & y_{\min} \leq \tilde{y}_j(\tilde{x}^q) \leq y_{\max} \\ 0, & \tilde{y}_j(\tilde{x}^q) \geq y_{\max} \end{cases} \quad \text{STB response} \tag{8c}$$

Utilizing the $\tilde{d}_j(\tilde{y}_j(\tilde{x}^q))$ functions, obtain the $\tilde{d}_j^{a,b,c,d,e}(\tilde{x}^q)$ values for each response j . The \tilde{U}_j (Eq. 9) and \tilde{L}_j (Eq. 10); upper and lower deviation values, respectively, are then estimated as:

$$\tilde{U}_j^{a,b,c,d,e} = \tilde{d}_j(y_j(\tilde{x}^j)) = \tilde{d}_{jj} = (\tilde{U}_j^a, \tilde{U}_j^b, \tilde{U}_j^c, \tilde{U}_j^d, \tilde{U}_j^e), \forall j \tag{9}$$

$$\tilde{L}_j^{a,b,c,d,e} = \text{Min} \{ \tilde{d}_j(y_j(\tilde{x}^{(1)})), \dots, \tilde{d}_j(y_j(\tilde{x}^{(Q)})) \} = (\tilde{L}_j^a, \tilde{L}_j^b, \tilde{L}_j^c, \tilde{L}_j^d, \tilde{L}_j^e), \forall j \tag{10}$$

Let $\tilde{D}_j(\tilde{y}_j(\tilde{x}^q))$ denotes the deviation function of $\tilde{y}_j(\tilde{x}^q)$ and is calculated as (Eq. 11):

$$\tilde{D}_j(\tilde{y}_j(\tilde{x}^q)) = \frac{\tilde{y}_j^e(\tilde{x}^q) - \tilde{y}_j^d(\tilde{x}^q)}{1 - \lambda}, \forall j \tag{11}$$

Then, calculate the $\tilde{D}_j(\tilde{y}_j(\tilde{x}^q))$ values for all $\tilde{y}_j(\tilde{x}^q)$. Finally, calculate the \tilde{P}_j (Eq. 12) and \tilde{Q}_j (Eq. 13) as follows:

$$\tilde{P}_j^{a,b,c,d,e} = \tilde{D}_j(\tilde{x}^j) = \tilde{D}_{jj} = (\tilde{P}_j^a, \tilde{P}_j^b, \tilde{P}_j^c, \tilde{P}_j^d, \tilde{P}_j^e), \forall j \tag{12}$$

$$\tilde{Q}_j^{a,b,c,d,e} = \text{Max} \{ \tilde{D}_j(\tilde{y}_j(\tilde{x}^{(1)})), \dots, \tilde{D}_j(\tilde{y}_j(\tilde{x}^{(Q)})) \} = (\tilde{Q}_j^a, \tilde{Q}_j^b, \tilde{Q}_j^c, \tilde{Q}_j^d, \tilde{Q}_j^e), \forall j \tag{13}$$

Step IV: Formulate the complete fuzzy optimization models.

The optimization model consists of two multiple objectives; to maximize desirability and minimize the deviation. That is,

$$\text{Max}\{\text{Desirability Function}\} = \text{Max}\{\tilde{d}_1(\tilde{y}_1(\tilde{x})), \dots, \tilde{d}_j(\tilde{y}_j(\tilde{x}))\}$$

$$\text{Min}\{\text{Deviation Function}\} = \text{Min}\{\tilde{D}_1(\tilde{y}_1(\tilde{x})), \dots, \tilde{D}_j(\tilde{y}_j(\tilde{x}))\}$$

s. t.

$$x \in [\text{Factor levels}]$$

To formulate the model with a single objective function, two fuzzy functions; $\tilde{S}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^j))$ and $\tilde{T}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^j))$, will be introduced which are calculated respectively using Eqs. 14 and 15.

$$\tilde{S}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^j)) = \begin{cases} 0, & \tilde{d}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^j)) \leq \tilde{L}_j^{a,b,c,d,e} \\ \frac{\tilde{d}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^j)) - \tilde{L}_j^{a,b,c,d,e}}{\tilde{U}_j^{a,b,c,d,e} - \tilde{L}_j^{a,b,c,d,e}}, & \tilde{L}_j^{a,b,c,d,e} \leq \tilde{d}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^j)) \leq \tilde{U}_j^{a,b,c,d,e} \forall j \\ 1, & \tilde{d}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^j)) \geq \tilde{U}_j^{a,b,c,d,e} \end{cases} \quad (14)$$

$$\tilde{T}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^j)) = \begin{cases} 1, & \tilde{D}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^j)) \leq \tilde{P}_j^{a,b,c,d,e} \\ \frac{\tilde{Q}_j^{a,b,c,d,e} - \tilde{D}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^j))}{\tilde{Q}_j^{a,b,c,d,e} - \tilde{P}_j^{a,b,c,d,e}}, & \tilde{P}_j^{a,b,c,d,e} \leq \tilde{D}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^j)) \leq \tilde{Q}_j^{a,b,c,d,e} \forall j \\ 0, & \tilde{D}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^j)) \geq \tilde{Q}_j^{a,b,c,d,e} \end{cases} \quad (15)$$

Let

$$\text{Min } \tilde{S}_j(\tilde{y}_j(\tilde{x}^j)) = \tilde{S} \quad (16)$$

$$\text{Min } \tilde{T}_j(\tilde{y}_j(\tilde{x}^j)) = \tilde{T} \quad (17)$$

Let w_1 and w_2 represent the assigned weights for desirability and robustness, which are usually chosen by decision maker based on cost and warranty. Then, formulate the final optimization model as:

$$\begin{aligned} \text{Max } & w_1 \tilde{S}^{a,b,c,d,e} + w_2 \tilde{T}^{a,b,c,d,e} \\ \text{s. t. } & \tilde{d}_j^{a,b,c,d,e}(x) - \tilde{S}^{a,b,c,d,e}(\tilde{U}_j^{a,b,c,d,e} - \tilde{L}_j^{a,b,c,d,e}) \geq \tilde{L}_j^{a,b,c,d,e} \quad , \forall j \\ & \tilde{D}_j^{a,b,c,d,e}(x) + \tilde{T}^{a,b,c,d,e}(\tilde{Q}_j^{a,b,c,d,e} - \tilde{P}_j^{a,b,c,d,e}) \leq \tilde{Q}_j^{a,b,c,d,e} \quad , \forall j \\ & 0 \leq \tilde{S}^{a,b,c,d,e} \leq 1 \\ & 0 \leq \tilde{T}^{a,b,c,d,e} \leq 1 \\ & x \in [\text{Factor levels}]. \end{aligned}$$

Solve the optimization model to identify the optimal factor levels for each of the fuzzy number elements; a, b, c, d , and e .

Step V: Apply the proposed optimization procedure on real case studies, and then analyze and discuss the optimization results. Finally, compare the anticipated improvements in the fuzzy quality characteristics by the proposed optimization procedure with those using other technique(s) in previous studies.

2.2 Experimental work

Several studies [13-18] were conducted to optimize process performance for a product's multiple quality characteristics. In this research, two processes were employed to illustrate the proposed optimization procedure and presented as follows:

Case study 1: Wire electro-discharge machining

Ramakrishnan and Karunamoorthy [13] optimized performance of wire electro-discharge machining (WEDM) of material removal rate (MRR, y_1) and surface roughness (SR, y_2) using artificial neural network (ANN) models and multi-response optimization technique. The MRR and SR the larger-the-better (LTB) and the smaller-the-better (STB) type quality characteristics, respectively. Four three-level controllable process factors were examined; the pulse on time (x_1), delay time (x_2), wire feed speed (x_3), and ignition current (x_4). The Taguchi's L_9 orthogonal array was utilized for experimental design as shown in Table 1. After randomization, each experiment was conducted. Experimentation was replicated twice. Finally, the MRR (mm^2/min) and SR (mm^2/min) values were measured. Let y_{i11} and y_{i12} denote the MRR values for the first and second replicate at the i -th experiment, respectively. Similarly, let y_{i21} and y_{i22} denote the SR values for the first and second replicate at the i -th experiment, respectively. Table 1 also displays the measured MRR and SR values for all experiments.

Table 1 Results of designed experiments for WEDM [13]

Exp. <i>i</i>	Control factor				MRR (mm ² /min)		SR(mm ² /min)	
	<i>x</i> ₁	<i>x</i> ₂	<i>x</i> ₃	<i>x</i> ₄	<i>y</i> _{i11}	<i>y</i> _{i12}	<i>y</i> _{i21}	<i>y</i> _{i22}
1	0.6	4	8	8	46	46	3.2	3.1
2	0.6	6	12	12	48	47	3.3	3.2
3	0.6	8	15	16	42	41	3.3	3.3
4	0.8	4	12	16	56	55	3.8	3.7
5	0.8	6	15	8	50	49	3.4	3.5
6	0.8	8	8	12	52	53	3.2	3.3
7	1.2	4	15	12	70	71	4.2	4
8	1.2	6	8	16	74	73	3.8	3.5
9	1.2	8	12	8	46	46	3.2	3.1

The multilinear regression models for MRR (*y*₁₁, *y*₁₂) and SR replicates, (*y*₂₁ and *y*₂₂) are expressed respectively as:

Regression model for response replicate	<i>R</i> ² _{adjusted}
<i>y</i> ₁₁ = -2.68 + 41.88 <i>x</i> ₁ + 4.26 <i>x</i> ₂ - 1.251 <i>x</i> ₃ + 3.82 <i>x</i> ₄ - 0.458 <i>x</i> ₂ ² + 0.119 <i>x</i> ₂ <i>x</i> ₃ - 0.166 <i>x</i> ₄ ²	99.48 %
<i>y</i> ₁₂ = -2.53 + 42.319 <i>x</i> ₁ + 4.21 <i>x</i> ₂ - 1.175 <i>x</i> ₃ + 3.76 <i>x</i> ₄ - 0.333 <i>x</i> ₂ ² + 0.1087 <i>x</i> ₂ <i>x</i> ₃ - 0.1934 <i>x</i> ₄ ²	99.99 %
<i>y</i> ₂₁ = 2 + 2.207 <i>x</i> ₁ - 0.04 <i>x</i> ₂ + 0.0233 <i>x</i> ₃ + 0.0284 <i>x</i> ₄ - 0.278 <i>x</i> ₁ ² - 0.1364 <i>x</i> ₁ <i>x</i> ₂ + 0.00417 <i>x</i> ₂ ²	97.43 %
<i>y</i> ₂₂ = 1.326 + 2.155 <i>x</i> ₁ - 0.0446 <i>x</i> ₂ + 0.0254 <i>x</i> ₃ + 0.0199 <i>x</i> ₄ - 0.2805 <i>x</i> ₁ ² - 0.169 <i>x</i> ₁ <i>x</i> ₂ + 0.0083 <i>x</i> ₂ ²	98.30 %

The fuzzy numbers of the intercept and regression coefficients are calculated and then the multiple regression, $\tilde{y}_1(\tilde{x})$, for *y*₁ is constructed as ($\lambda = 0.15$):

$$\tilde{y}_1(\tilde{x}) = (-2.9, -2.71, -2.68, -2.65, -2.46) + (41.27, 41.79, 41.88, 41.97, 42.49)\tilde{x}_1 + (4.18, 4.24, 4.26, 4.27, 4.33)\tilde{x}_2 + (-1.26, -1.22, -1.21, -1.2, -1.15)\tilde{x}_3 + (3.73, 3.8, 3.82, 3.83, 3.9)\tilde{x}_4 + (-0.63, -0.48, -0.45, -0.43, -0.28)\tilde{x}_2^2 + (0.1, 0.11, 0.12, 0.121, 0.13)\tilde{x}_2\tilde{x}_3 + (-0.2, -0.17, -0.166, -0.16, -0.13)\tilde{x}_2\tilde{x}_4$$

The multiple regression, $\tilde{y}_2(\tilde{x})$, for *y*₂ is constructed in a similar manner as:

$$\tilde{y}_2(\tilde{x}) = (0.37, 1.18, 1.32, 1.46, 2.27) + (2.08, 2.14, 2.15, 2.16, 2.22)\tilde{x}_1 + (-0.051, -0.045, -0.044, -0.043, -0.038)\tilde{x}_2 + (0.022, 0.025, 0.0254, 0.026, 0.028)\tilde{x}_3 + (0.007, 0.018, 0.019, 0.021, 0.031)\tilde{x}_4 + (-0.284, -0.281, -0.28, -0.279, -0.276)\tilde{x}_1^2 + (-0.21, -0.17, -0.169, -0.162, -0.122)\tilde{x}_1\tilde{x}_2 + (0.002, 0.007, 0.008, 0.009, 0.014)\tilde{x}_2^2$$

Utilizing the formulations in Steps I and II, the optimization model is developed and then solved to determine the combination of optimal factor settings for each replicate of the two responses. The obtained optimization results for each replicate are displayed for all response replicates in Table 2. The optimization models for \tilde{y}_{12} , \tilde{y}_{21} , and \tilde{y}_{22} , are formulated and then solved in a similar manner. The results of optimal factor settings are also shown in Table 2.

The results shown in Table 2 are utilized in Steps III and IV to generate the pentagon fuzzy numbers for optimal factor settings; $\tilde{x}^{(1)}$ and $\tilde{x}^{(2)}$, of \tilde{y}_1 and \tilde{y}_2 , respectively. The obtained $\tilde{y}_j(\tilde{x}^q)$ values are displayed in Table 3.

The $\tilde{d}_1(\tilde{y}_1(\tilde{x}^q))$, $\tilde{d}_2(\tilde{y}_2(\tilde{x}^q))$, $\tilde{D}_1(\tilde{y}_1(\tilde{x}^q))$, and $\tilde{D}_2(\tilde{y}_2(\tilde{x}^q))$ values are then calculated for both responses and the results are shown in Table 4.

Table 2 Optimal results for WEDM

Optimal level	Response replicate			
	\tilde{y}_{11}	\tilde{y}_{12}	\tilde{y}_{21}	\tilde{y}_{22}
<i>x</i> ₁ [*]	0.9341	1.0254	0.6000	1.1345
<i>x</i> ₂ [*]	4.0000	4.0000	4.0000	8.0000
<i>x</i> ₃ [*]	8.0000	13.6120	8.0000	8.0000
<i>x</i> ₄ [*]	8.0000	8.0000	8.0000	8.0000

Table 3 The calculated $\tilde{y}_j(\tilde{x}^q)$ values for WEDM

Optimal combination	$\tilde{y}_1(\tilde{x}^q)$	$\tilde{y}_2(\tilde{x}^q)$
$\tilde{x}^{(1)}$	(52.46, 58.8, 60, 61.2, 68.5)	(1.29, 2.65, 2.89, 3.13, 4.54)
$\tilde{x}^{(2)}$	(33.5, 54.65, 58.61, 62.7, 89.91)	(1.09, 2.26, 2.49, 2.74, 4.53)

Table 4 The calculated matrix of $\tilde{d}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^q))$ and $\tilde{D}_j^{a,b,c,d,e}(\tilde{y}_j(\tilde{x}^q))$ of MMR and SR

	$\tilde{d}_1(\tilde{y}_1(\tilde{x}^q))$	$\tilde{d}_2(\tilde{y}_2(\tilde{x}^q))$	$\tilde{D}_1(\tilde{y}_1(\tilde{x}^g))$	$\tilde{D}_2(\tilde{y}_2(\tilde{x}^g))$
$\tilde{x}^{(1)}$	(0.098, 0.352, 0.399, 0.448, 0.7402)	(0.012, 0.317, 0.371, 0.423, 0.717)	(7.746, 8.164, 8.238, 8.312, 8.73)	(1.432, 1.457, 1.461, 1.466, 1.491)
$\tilde{x}^{(2)}$	(0, 0.186, 0.344, 0.508, 1)	(0.014, 0.402, 0.457, 0.508, 0.762)	(6.41, 10.78, 11.762, 12.81, 19.94)	(1.263, 1.567, 1.635, 1.708, 2.203)

The optimization models *a* to *e* were formulated for WEDM and then solved. For illustration, the Model *a* is formulated as:

$$\begin{aligned}
 & \text{Max } 0.5 \times S^a + 0.5 \times T^a \\
 & \text{s. j.} \\
 & -2.11 + 1.65x_1^a + 0.167x_2^a - 0.05x_3^a + 0.149x_4^a - 0.025x_2^ax_2^a + 0.0041x_2^ax_3^a - 0.0081x_4^ax_4^a - 0.0985S^a \geq 0 \\
 & 0.504 - 0.484x_1^a + 0.0082x_2^a - 0.0062x_3^a - 0.0069x_4^a + 0.0602x_1^ax_1^a + 0.026x_1^ax_2^a - 0.0031x_2^ax_2^a - 0.0026S^a \geq 0.01208 \\
 & 0.219 + 0.6144x_1^a + 0.0707x_2^a + 0.0537x_3^a + 0.084x_4^a + 0.1765x_2^ax_2^a + 0.015x_2^ax_3^a + 0.0374x_4^ax_4^a + 0T^a \leq 7.7462 \\
 & 0.953 + 0.0735x_1^a + 0.0065x_2^a + 0.003x_3^a + 0.012x_4^a + 0.0035x_1^ax_1^a + 0.0468x_1^ax_2^a + 0.0058x_2^ax_2^a + 0.1691T^a \leq 1.4323 \\
 & 0 \leq S^a \leq 1 \\
 & 0 \leq T^a \leq 1
 \end{aligned}$$

The optimization models *b* to *e* are formulated and then solved in a similar manner.

Case study 2: Sputtering process of gallium-doped zinc oxide GZO films

Chen *et al.* [14] optimized five process controllable factors; R.F. power (x_1), sputtering pressure (x_2), deposition time (x_3), substrate temperature (x_4), and post-annealing temperature (x_5) of GZO films deposited on polyethylene terephthalate substrates by magnetron sputtering using the Taguchi method. Three important quality responses were considered including: deposition rate (DR, y_1) electrical resistivity (ER, y_2) and optical transmittance (OT, y_3), which are LTB, STB, and LTB type quality characteristics, respectively. The Taguchi's L_{18} orthogonal array shown in Table 5 was employed to provide experimental layout, where each experiment was repeated twice. Let y_{i1r} , y_{i2r} , and y_{i3r} denote the r -th replicate ($r = 1$ or 2) of DR, ER, and OT responses at experiment i ; $i = 1, \dots, 18$, respectively. Table 5 also displays the experimental data of y_{i1r} , y_{i2r} , and y_{i3r} for the sputtering process.

Following the proposed procedure, the optimization models were constructed to determine the combination of optimal factor settings that optimizes each response replicate. Table 6 lists the obtained optimal factor settings for each replicate for all responses.

Utilizing the results in Table 6, the optimization model *a* to *e* were formulated for the sputtering process and then the combination of optimal fuzzy process factors were determined.

Table 5 Experimental data for sputtering process [14]

Exp. <i>i</i>	Factor					DR		ER		OT	
	x_1	x_2	x_3	x_4	x_5	y_{i11}	y_{i12}	y_{i21}	y_{i22}	y_{i31}	y_{i32}
1	50	0.13	30	25	0	4.5	4.7	14.9	15.3	88.4	88.4
2	50	0.67	60	50	100	5.6	5.6	9.8	9.7	87.7	87.7
3	50	1.33	90	100	200	5.0	4.9	7.9	7.8	88.1	88.1
4	100	0.13	30	50	100	9.6	9.3	5.4	5.6	89.2	89.3
5	100	0.67	60	100	200	11.1	11.3	4.6	4.3	87.1	87.0
6	100	1.33	90	25	0	10.0	10.0	6.5	6.6	84.7	84.7
7	200	0.13	60	25	200	19.9	20.2	1.6	1.7	86.6	86.6
8	200	0.67	90	50	0	21.6	21.6	1.9	2.0	82.3	82.4
9	200	1.33	30	100	100	20.9	20.9	1.8	1.6	85.6	85.3
10	50	0.13	90	100	100	4.8	4.6	7.3	7.0	87.6	87.6
11	50	0.67	30	25	200	5.0	4.9	6.9	7.1	89.1	89.1
12	50	1.33	60	50	0	4.9	4.8	7.8	7.7	87.4	87.4
13	100	0.13	60	100	0	9.7	9.7	6.1	5.9	87.0	87.0
14	100	0.67	90	25	100	11.1	11.6	6.0	5.8	83.7	83.7
15	100	1.33	30	50	200	10.7	10.8	5.5	5.7	88.4	88.3
16	200	0.13	90	50	200	19.5	19.4	1.0	1.1	83.1	83.1
17	200	0.67	30	100	0	22.1	22.0	1.2	1.3	85.7	85.7
18	200	1.33	60	25	100	20.5	20.5	1.4	1.3	83.9	83.7

Table 6 Optimization results for sputtering process for each response replicate

Factors	y_{11}	y_{12}	y_{21}	y_{22}	y_{31}	y_{32}
x_1^*	92.60	94.03	200.00	200.00	50.00	50.00
x_2^*	1.33	1.33	1.33	1.33	1.31	0.13
x_3^*	30.00	30.00	30.00	30.00	30.00	30.00
x_4^*	25.00	25.00	25.00	25.00	25.50	25.00
x_5^*	200.00	1.24	1.22	4.71	1.24	1.24

3. Results and discussion

In Step V, two case studies were utilized to illustrate the proposed procedure. The optimal results for each case study are discussed and compared with those obtained by previously used techniques in the following subsections.

3.1 Optimal wire electrical-discharge machining

Using Lingo 11 package, the combinations of the fuzzy optimal factor settings for each of models a , b , c , d and e of the Wire electrical-discharge machining were obtained and are then listed in Table 7. For example, the obtained combination of optimal factor settings (x_1^* , x_2^* , x_3^* , x_4^*) by solving model a is (0.9019, 4.00, 8.00, 8.331).

Finally, the \tilde{y}_1 and \tilde{y}_2 values of MRR and SR, respectively, are calculated at all combinations of optimal factor settings and found to be (51.05, 56.96, 59.22, 61.21, 70.02) and (1.31, 2.48, 2.72, 2.89, 4.16) mm²/min, respectively. The results reveal the effect of process variability on MRR (LTB type) and SR (STB type). For illustration, when the combination of optimal process factor settings by model a is selected, the MRR and SR are 51.05 and 1.31, respectively. However, when the optimal combination of factor settings by model e is chosen, the MRR and SR are 70.02 and 4.16, respectively. Clearly, model a provides the largest improvement in SR, whereas model e provides the largest improvement in MRR. Consequently, process engineering can select and then control the proper combination of optimal settings from models a to e that satisfies product requirements.

Table 7 Fuzzy optimal factor settings for WEDM

Optimal level	Model a	Model b	Model c	Model d	Model e
\tilde{x}_1^*	0.9019	0.8809	0.9045	0.9067	0.8810
\tilde{x}_2^*	4.0000	4.1870	4.2220	4.6920	4.9460
\tilde{x}_3^*	8.0000	8.0000	8.0000	8.0000	8.0000
\tilde{x}_4^*	8.3310	8.0000	8.0000	8.0000	8.8750

3.2 Optimal sputtering process

Solving models a to e of the sputtering process, the combination of optimal factor settings (x_1^* , x_2^* , x_3^* , x_4^*) were found determined and displayed in Table 8.

Then, the \tilde{y}_1 , \tilde{y}_2 , and \tilde{y}_3 values for DR (nm/min), ER (10^{-4} Ω cm), and OT (%), respectively, were calculated at the optimal factor settings of models a , b , c , d and e and found to be (11.895, 12.282, 14.886, 18.071, 21.946), (2.848, 3.942, 4.098, 4.784, 5.19) and (88.218, 88.276, 88.395, 88.626, 89.174), respectively. Obviously, the combinations of optimal factor settings vary due to the observed fuzziness in the DR, ER, and OT. For example, the DR value at optimal setting of model a is 11.895, while it is 21.946 using optimal setting of model e . Note that the optimal setting of factor x_4 , x_4^* , changes from 62.852 to 25. Similarly, optimal setting of factor x_1 , x_1^* , changes from 127.605 to 195.153. Although the optimal settings of model e results in improving the DR (LTB type), but it worsens the ER (STB type). Consequently, process engineers should decide and control the settings of these two factors to avoid the negative impacts of their variations on quality characteristics.

Table 8 Fuzzy optimal factor settings for sputtering process

Factor	Model <i>a</i>	Model <i>b</i>	Model <i>c</i>	Model <i>d</i>	Model <i>e</i>
x_1^*	127.605	129.091	143.363	161.715	195.153
x_2^*	0.130	0.130	0.130	0.130	0.130
x_3^*	30.00	30.000	30.000	30.000	30.000
x_4^*	62.852	63.852	69.967	60.160	25.000
x_5^*	0.000	0.000	0.000	0.000	0.000

3.3 Comparison of the results

For WEDM process, the comparison between the optimization results by ANN [13] and these obtained by the proposed procedure is made in Table 9, where it is noticed that at the initial (ANN) process factor settings of WEDM process, the values of MRR (mm²/min) and SR (mm²/min) were 64 (71) and 3.48 (3.198), respectively. However, these approaches fail to consider the effects of process variations/fuzziness on quality responses. Using the proposed procedure, the MRR and SR values of (51.05, 56.96, 59.22, 61.21, 70.02) and (1.31, 2.48, 2.72, 2.89, 4.16), respectively. This result clearly reveals the existence of fuzziness between the observed measurements of each response replicates, which is ignored when optimizing process settings by ANN, thereby misleading process engineers. Moreover, the proposed procedure provides better understanding on how to choose the optimal factor settings that satisfy both quality responses. In other words, if the SR (mm²/min) is more important than the MRR, the minimal value that can be reached by ANN technique is 3.198. However, utilizing the optimal settings of model *a* from the proposed approach results in reducing the values of SR and MRR to 1.31 and 51.05, respectively. Further, the ANN technique requires defining initial parameter settings of learning and prediction stages which are usually unknown, which the uncertainty in the decision making process about optimal factor settings.

Table 10 displays the comparison between the DR, ER, and OT values at the initial and the optimal factor settings of the sputtering process using the Grey-Taguchi [14] and the proposed approach. It is found that the DR (nm/min), ER (10⁻⁴ Ω cm), and OT (%) values at initial (Grey-Taguchi) factor settings were 21.033 (20.922), 11.9 (8.627) and 86.627 % (90.00 %). While, their corresponding values by using the proposed procedure are (11.895, 12.282, 14.886, 18.071, 21.946), (2.848, 3.942, 4.098, 4.784, 5.19), and (88.218, 88.276, 88.395, 88.626, 89.174), respectively. Obviously, settings process factors of sputtering process at those obtained by model *e* of the proposed approach leads to achieve significant improvement; that is, DR, ER, and OT values of 21.946, 5.19, and 89.174, respectively, in the three quality responses when compared to those at the initial and the Grey-Taguchi method. Further, the Grey-Taguchi method is a non-parametric approach that cannot guarantee optimal factor settings and failed to consider the effect of fuzziness on the three quality characteristics.

Table 9 MRR and SR at the combination of optimal fuzzy factor settings of WEDM

Response	Initial setting	ANN [13]	Proposed models
MRR (mm ² /min)	64	71	(51.05, 56.96, 59.22, 61.21, 70.02)
SR (mm ² /min)	3.48	3.198	(1.31, 2.48, 2.72, 2.9, 4.16)

Table 10 MRR and SR at the combination of optimal fuzzy factor settings of sputtering process

Response	Initial setting	Grey-Taguchi [14]	Proposed models
DR (nm/min, LTB)	21.033	20.922	(11.895, 12.282, 14.886, 18.071, 21.946)
ER (10 ⁻⁴ Ω cm, STB)	11.9	8.627	(2.848, 3.942, 4.098, 4.784, 5.19)
OT (% , LTB)	86.148	90	(88.218, 88.276, 88.395, 88.626, 89.174)

In summary, the proposed procedure has the following advantages:

- effectiveness in dealing with fuzziness through the use of pentagon fuzzy regression models;
- depiction of the relationship between quality response and process factors, which helps process engineers in identifying the critical process factors and determining the required adjustments in process levels to increase the anticipated improvements in desirable quality responses;

- guaranteeing optimality of factor levels, and
- conveying valuable information to process engineers in understanding the impact of fuzziness on process performance and then guiding them to take proper improvement actions.

4. Conclusion

This research proposed an efficient procedure for optimizing process performance for fuzzy multiple quality responses utilizing pentagon regression modelling. Initially, the optimal factor settings were determined for each replicate of a quality response. Then, the optimal factor settings for the replicates of each quality response were combined to construct a pentagon fuzzy regression model for this quality response. The combinations of fuzzy responses, desirability, and deviations values were constructed and utilized in formulating an optimization model, which was then solved for each element of the pentagon fuzzy number to decide optimal factor settings for multiple quality responses concurrently. Two case studies were employed for illustration, where the results of both cases revealed that the proposed procedure efficiently dealt with fuzziness problem in quality responses. In conclusion, the developed optimization procedure provided valuable assistance to process engineering when optimizing process performance for fuzzy multiple quality responses.

Conflicts of interest

The authors declare no conflict of interest.

References

- [1] Athreya, S., Venkatesh, Y.D. (2012). Application of Taguchi method for optimization of process parameters in improving the surface roughness of lathe facing operation, *International Refereed Journal of Engineering and Science*, Vol. 1, No. 3, 13-19.
- [2] Liao, H.-C., Chen, Y.-K. (2002). Optimizing multi-response problem in the Taguchi method by DEA based ranking method, *International Journal of Quality & Reliability Management*, Vol. 19, No. 7, 825-837, doi: [10.1108/02656710210434766](https://doi.org/10.1108/02656710210434766).
- [3] Lin, C.L. (2004). Use of the Taguchi method and grey relational analysis to optimize turning operations with multiple performance characteristics, *Materials and Manufacturing Processes*, Vol. 19, No. 2, 209-220, doi: [10.1081/AMP-120029852](https://doi.org/10.1081/AMP-120029852).
- [4] Hocine, A., Kouaissah, N., Bettahar, S., Benbouziane, M. (2018). Optimizing renewable energy portfolios under uncertainty: A multi-segment fuzzy goal programming approach, *Renewable Energy*, Vol. 129, Part A, 540-552, doi: [10.1016/j.renene.2018.06.013](https://doi.org/10.1016/j.renene.2018.06.013).
- [5] Komsiyah, S.; Meiliana; Centika, H.E. (2018). A fuzzy goal programming model for production planning in furniture company, *Procedia Computer Science*, Vol. 135, 544-552, doi: [10.1016/j.procs.2018.08.207](https://doi.org/10.1016/j.procs.2018.08.207).
- [6] Mirzaee, H., Naderi, B., Pasandideh, S.H.R. (2018). A preemptive fuzzy goal programming model for generalized supplier selection and order allocation with incremental discount, *Computers & Industrial Engineering*, Vol. 122, 292-302, doi: [10.1016/j.cie.2018.05.042](https://doi.org/10.1016/j.cie.2018.05.042).
- [7] Johnson, D.B., Bogle, I.D.L. (2019). A quantitative risk analysis approach to a process sequence under uncertainty – A case study, *Computers & Chemical Engineering*, Vol. 126, 1-21, doi: [10.1016/j.compchemeng.2019.03.039](https://doi.org/10.1016/j.compchemeng.2019.03.039).
- [8] Al-Refaie, A., Musallam, A. (2018). Using mixed goal programming to optimize performance of extrusion process for multiple responses of irrigation pipes, *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, Vol. 233, No. 2, 412-424, doi: [10.1177/0954408918781624](https://doi.org/10.1177/0954408918781624).
- [9] Yaghoobi, M.A., Jones, D.F., Tamiz, M. (2008). Weighted additive models for solving fuzzy goal programming problems, *Asia-Pacific Journal of Operational Research*, Vol. 25, No. 5, 715-733, doi: [10.1142/S0217595908001973](https://doi.org/10.1142/S0217595908001973).
- [10] Gupta, A., Singh, H., Aggarwal, A. (2011). Taguchi-fuzzy multi output optimization (MOO) in high speed CNC turning of AISI P-20 tool steel, *Expert Systems with Applications*, Vol. 38, No. 6, 6822-6828, doi: [10.1016/j.eswa.2010.12.057](https://doi.org/10.1016/j.eswa.2010.12.057).
- [11] Dalman, H. (2016). An interactive fuzzy goal programming algorithm to solve decentralized bi-level multiobjective fractional programming problem, arXiv:1606.00927, math.OC, from <https://arxiv.org/abs/1606.00927>, accessed August 7, 2020.
- [12] Al-Refaie, A., Bani Domi, G., Abdullah, R. (2019). A fuzzy goal programming-regression approach to optimize process performance of multiple responses under uncertainty, *International Journal of Management Science and Engineering Management*, Vol. 14, No. 1, 20-32, doi: [10.1080/17509653.2018.1467802](https://doi.org/10.1080/17509653.2018.1467802).

- [13] Ramakrishnan, R., Karunamoorthy, L. (2008). Modeling and multi-response optimization of Inconel 718 on machining of CNC WEDM process, *Journal of Materials Processing Technology*, Vol. 207, No. 1-3, 343-349, doi: [10.1016/j.jmatprotec.2008.06.040](https://doi.org/10.1016/j.jmatprotec.2008.06.040).
- [14] Chen, C.-C., Tsao, C.-C., Lin, Y.-C., Hsu, C.-Y. (2010). Optimization of the sputtering process parameters of GZO films using the Grey-Taguchi method, *Ceramics International*, Vol. 36, No. 3, 979-988, doi: [10.1016/j.ceramint.2009.11.019](https://doi.org/10.1016/j.ceramint.2009.11.019).
- [15] Tahir, W., Jahanzaib, M., Raza, A. (2019). Effect of process parameters on cutting speed of wire EDM process in machining HSLA steel with cryogenic treated brass wire, *Advances in Production Engineering & Management*, Vol. 14, No. 2, 143-152, doi: [10.14743/apem2019.2.317](https://doi.org/10.14743/apem2019.2.317).
- [16] Evran, S., Mutlu, B., Kurt, M. (2019). Cutting rate and surface characteristic analysis in CNC wire electrical discharge machining of aluminium bronze, *Technical Gazette – Tehnički Vjesnik*, Vol. 26, No. 5, 1228-1233, doi: [10.17559/TV-20180203185742](https://doi.org/10.17559/TV-20180203185742).
- [17] Bhuyan, R.K., Routara, B.C., Parida, A.K. (2015). Using entropy weight, OEC and fuzzy logic for optimizing the parameters during EDM of Al-24 % SiCP MMC, *Advances in Production Engineering & Management*, Vol. 10, No. 4, 217-227, doi: [10.14743/apem2015.4.204](https://doi.org/10.14743/apem2015.4.204).
- [18] Rao, P.S., Ramji, K., Satyanarayana, B. (2014). Experimental investigation and optimization of wire EDM parameters for surface roughness, MRR and white layer in machining of aluminium alloy, *Procedia Materials Science*, Vol. 5, 2197-2206, doi: [10.1016/j.mspro.2014.07.426](https://doi.org/10.1016/j.mspro.2014.07.426).

An integrated CNC system for chatter suppression in turning

Jasiewicz, M.^a, Miądlicki, K.^{a,*}

^aFaculty of Mechanical Engineering and Mechatronics, West Pomeranian University of Technology in Szczecin, Szczecin, Poland

ABSTRACT

Self-excited chatter vibrations are one of factors affecting the reduction of cutting efficiency, especially while machining highly compliant machine parts. Their occurrence can be limited by the proper technological parameters selection. These parameters can be determined by analysing the cutting process stability, which requires knowledge of the machine-tool-workpiece system dynamic properties. Normally, these properties are determined experimentally, which is troublesome in industrial practice. This article presents a method in which dynamic properties are calculated by single-board computer integrated with a Computer Numerical Control (CNC) system, with no need to carry out additional experimental tests. It is possible with the receptance coupling approach which allows for obtaining the workpiece geometry by analyzing the machining program and then determining the machine tool – workpiece system dynamic properties. These properties are the input to the presented algorithm that facilitates the selection of cutting parameters enabling stable turning of highly compliant machine parts. The presented system is dedicated to turning but can also be adapted to determine the stability of milling with flexible tools.

© 2020 CPE, University of Maribor. All rights reserved.

ARTICLE INFO

Keywords:

Computer numerical control (CNC);
Machining chatter;
Vibrations;
Stability analysis;
Machine-tool spindle;
Cutting parameters;
Turning

*Corresponding author:

karol.miadlicki@zut.edu.pl
(Miądlicki, K.)

Article history:

Received 7 September 2020
Revised 6 October 2020
Accepted 9 October 2020

1. Introduction

Research works on the suppression of self-excited chatter vibrations that were initiated by Tlustý [1] and Tobias [2] who found that vibration during machining results from the loss of process stability. The chatter suppression is still the subject of scientific research conducted around the world, which still does not provide a universally effective solution of this problem. Performing the cutting with chatter vibration negatively affects the quality of the machined surface, shortens the tool life and can ultimately lead to machine tool failure [3]. Avoiding the chatter vibrations is a particularly challenging task while machining flexible parts (e.g. slender shafts turning), as presented by Powalka et al. in [4] or Kaliński et al. in [5]. A detailed review of the most important chatter suppression techniques and chatter vibration research was presented by M. Siddhpura et al. in [6] and J. Munoa et al. in [7].

Basically, two approaches concerning chatter suppression can be distinguished: introduction of process modification to maintain its stability or search for a range of stability within the existing process [8]. The first group consists of hardware solutions and may refer to solutions introduced at the stage of machine tool construction or implementation of improvements to the existing machine tools. An innovative approach to the construction of machine tools in terms of providing higher resistance to vibration during machining was presented by Dunaj et al. in [9–11]. They proposed the construction of the machine tool body as a welded steel-polymer concen-

trate frame, which significantly increases the vibration damping capacity. Another solution for increasing the vibration damping capacity is the introduction of additional damping elements to the structure such as epoxy coating of the linear guide system foundation for the milling machine proposed by Powalka et al. [12] or the introduction of additively manufactured vibration eliminators presented by Dunaj et al. in [13]. Process modification at a later stage that can be implemented on an existing machine tool is the use of active vibration eliminators, as proposed by Parus et al. in [14] or Brecher et al. in [15]. The second group of methods, consisting of searching for areas of machining stability, can be defined as a software solution. Ensuring cutting stability is done through a selection of machining parameters (spindle rotational speed and cutting depth) for which the chatter vibrations do not occur. The elementary method for selecting these parameters is stability lobes, presenting the cutting depths at which chatter occurs as a function of spindle rotational speed [16-18]. The stability lobes are determined based on the dynamic properties of using a machine tool-workpiece system. These properties, like frequency response function (FRFs) can be determined in numerous ways. The experimental methods assume carrying out impulse tests, where the accelerometers measure the system's response excited with a modal hammer [19]. However, this solution is dedicated mainly to research units with highly specialized measuring equipment and qualified staff to perform the experiment. An alternative to experimental measurements is the model-based approach, which consists in determining the machine tool-workpiece system FRFs using a finite element method [20-22]. Such methods, however, remain computationally complex, and achieving good results requires expensive software and highly experienced staff. A solution that incorporates the advantages of an experimental and model approach is the use of the receptance coupling [23]. The dynamic properties of the machine tool are determined experimentally and the compliant part of the system (a tool or a workpiece) is modeled analytically, as presented for milling by Park et al. in [24] or for turning by Jasiewicz et al. in [25].

The operation of CNC machine tools requires the operator experience in programming as well as knowledge of technological issues. Manufacturers of the most common CNC control systems provide solutions using a graphic interface, significantly simplifying programming in G-code (Shopturn/ShopMill by Siemens or Manual Guide by Fanuc). CNC control systems are also often modified by machine tool manufacturers by introducing their own human-machine interfaces (e.g., Celos by DMG Mori or Mazatrol by Mazak). This introduces the possibility of using different CNC systems in the same machine tool model while maintaining a consistent graphic design of the operator panels. Furthermore, touch screens, 3D graphics or gesture control are visually attractive and increase the comfort of using the machine. However, despite the improvement in terms of machine tool operation and programming, the issue of optimal technological parameters selection remains without additional support. In some cases, an arbitrary selection of the parameters proposed by the tool manufacturer will be sufficient. However, when machining compliant parts where there is a risk of vibration, additional support in this area can be particularly helpful. One of the solutions offered by machine tool manufacturers in the Machining Navi system by Okuma. Based on the acoustic analysis of the process, in case of vibrations during machining, it proposes to change the spindle speed. However, as in all commercial solutions, the operation of the system is a trade secret and the manufacturer does not provide technical details for the solution. Machining optimization systems that can be integrated into a machine tool have been the subject of research for many research teams. Passive systems for chatter suppression are based only on software solutions. Intelligent optimization of machining parameters for turning using the evolutionary algorithms was presented by Mia et al. in [26]. A solution in which computer numerical control parameters are optimized using fuzzy logic was proposed by Chiu et al. in [27]. Jasiewicz et al. [28] proposed a parameter selection assistant for turning integrated into the CNC system. Sun et al. developed in [29] a chatter detection algorithm based on variable-scale wavelet packet entropy. Active methods for chatter suppression are also very popular. Apart from software solutions, they use additional devices integrated with the CNC machine system. An active control approach using two piezoelectric actuators and an adaptive neural-network-based controller was proposed by Liu et al. in [30]. Another active control method consisting of an adaptive sliding-mode controller and a displacement field recon-

struction method for chatter suppression was developed by Ma et.al. in [31]. Chen et. al. in [32] used an in-house designed magnetic actuator for an active damping method of boring bars. The popularity of integrating passive and active systems solutions into CNC systems will grow due to increasing computational power of CNC systems and introducing build-in sensors in the machine-tools. This creates the possibility to continuous monitoring and impact on the machining process. Another difficulty in developing a universal solution is the large variety of materials and geometry of workpieces. In [33] highlights some factors which affect the formation of a protective built-up layer (BUL) on the rake face of the cutting tool when cutting magnesium alloys. In other work [34] the authors focus on the problem of control of the accuracy of forming elastic-deformable shafts with low rigidity. The paper focuses on analyzing fundamental factors affecting the accuracy of machining of low-rigidity shafts like stiffness, the geometry of the cutting tool, lathe temperature, degree of cutting tool wear, cutting tool strength, lubrication-cooling fluid, and machining parameters are presented.

This article presents the concept of a low-cost system supporting the selection of turning parameters on a lathe with a FANUC CNC system. The first part concerns the issues of ensuring machining stability and the procedure for determining the FRFs of the machine tool-workpiece system using the receptance coupling approach. The second part presents the idea of a system supporting the selection of machining parameters, and subsequently, the implementation of the assistant for a CNC lathe equipped with a FANUC control system. The summary presents the results of an experimental study as well as conclusions.

2. Materials and methods

This subsection presents issues related to the problems of self-excited chatter vibrations occurring during machining, and a method for determining the stability of machining using stability lobes. Moreover, the receptance coupling procedure is described, which allows for determining the machine tool-workpiece system's frequency response function, obligatory for evaluating the stability lobes.

2.1 Background on turning stability

The reason for the occurrence of vibration during machining is the loss of cutting process stability. Most cutting operations, including turning, involve machining so that the tool tip moves over a previously machined surface. If one of the vibration modes of the machine tool system (usually the dominant one, the first mode) is excited by cutting forces, a wavy trace is formed on the machined surface. In the case of turning, the edge of the tool will meet this trace at the next rotation, which leads to a temporary change of the chip thickness and hence, a temporary increase of cutting and as a result, the formation of a new wave (trace). This phenomenon is known as regenerative chatter. Research concerning this phenomenon carried out over the years shows that ensuring the stability of machining depends on the appropriate selection of the depth of cut. The limiting depth of cut a_{lim} for turning can be provided as:

$$a_{lim} = -\frac{1}{2K_r Re(G(j\omega))} \quad (1)$$

where: K_r – cutting force coefficient, $Re(G(j\omega))$ – real part of the FRF $G(j\omega)$ determined for the compliant part of the machine tool system.

For slender shaft turning the FRF is evaluated for the workpiece, while for a boring process, where tools usually have long overhangs, FRF is determined at the tool tip. From Eq. 1 it can be seen that to determine the cutting depth limit (which should be positive) negative values of the real part of the FRF are required, as shown in Fig. 1a. The part marked "negative real" is directly used to determine stability lobes, by converting to rotational speeds using Eq. 2 and then replicating for subsequent positive integers k as shown in Fig. 1b.

$$N_c = \frac{60 \cdot f_c}{k}, \quad \text{for } k = 1, 2, \dots, n \quad (2)$$

Variable f_c is the chatter frequency and k is a positive integer (representing the lobe number).

The stability lobes separate the stable and unstable machining area. Therefore, by selecting the appropriate spindle speed and cutting depth, chatter vibration can be avoided. The stability lobes presented in Fig. 1b require the FRF of the machine tool system and cutting force coefficient that describes the interaction between the workpiece (it is material) and the tool. In order to determine a stable cutting depth for specific spindle speed, the coefficients are evaluated experimentally for a specific set of the tool and the workpiece material. Omitting the influence of the coefficient does not allow for specifying the cutting depth limit, however it is still suitable for the selection of spindle speed. Although this solution does not guarantee full effectiveness in ensuring machining stability, the complex experimental procedure is omitted, making use of stability lobes more accessible for less advanced machining systems.

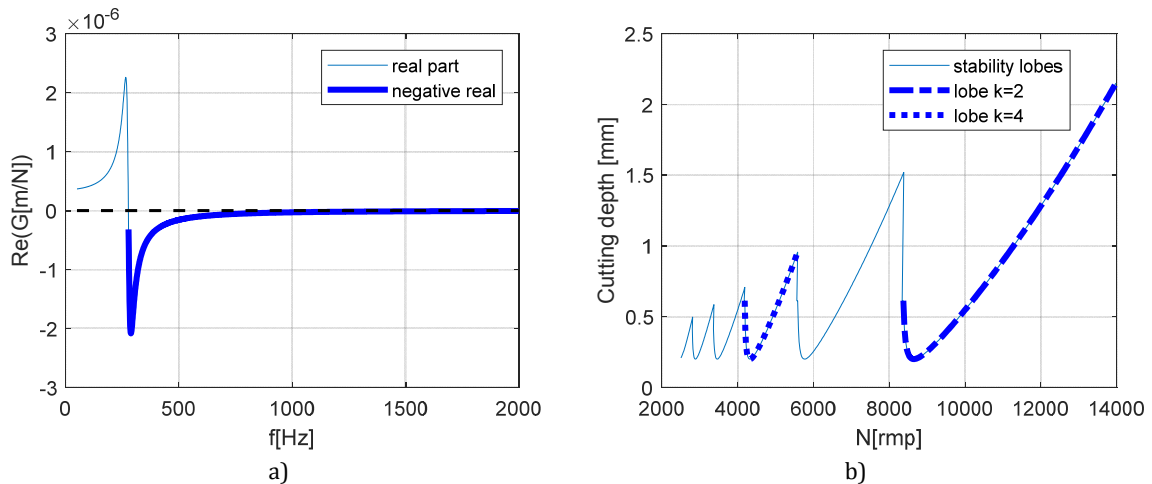


Fig. 1 a) Real part FRF of the machine tool system; b) Graphic representation of determining the stability lobes

2.2 Receptance coupling

The frequency response functions of the machine tool-workpiece system, needed for the stability lobes calculation, can be determined using the receptance coupling approach (RCA). To perform the synthesis of the dynamic properties of the coupled system, the FRFs for the selected points of its subassemblies are required. The developed procedure applies to the case in which slender shafts are machined. Therefore, this system consists of a lathe spindle with a three-jaw chuck (subassembly "1") and a workpiece, i.e. a rod (subassembly "2"). The components and their local coordinates are presented in Fig. 2a and the coupled system in Fig. 2b.

The modelled system consists of three points: "1" - point of the fixture spindle three-jaw chuck - workpiece, "2" a tool point which can be arbitrarily oriented along the axis of the workpiece, "3" point at the end of the workpiece. For turning, dynamic properties in the x direction have a major impact on chatter vibrations, therefore for the receptance coupling procedure the other directions have been neglected.

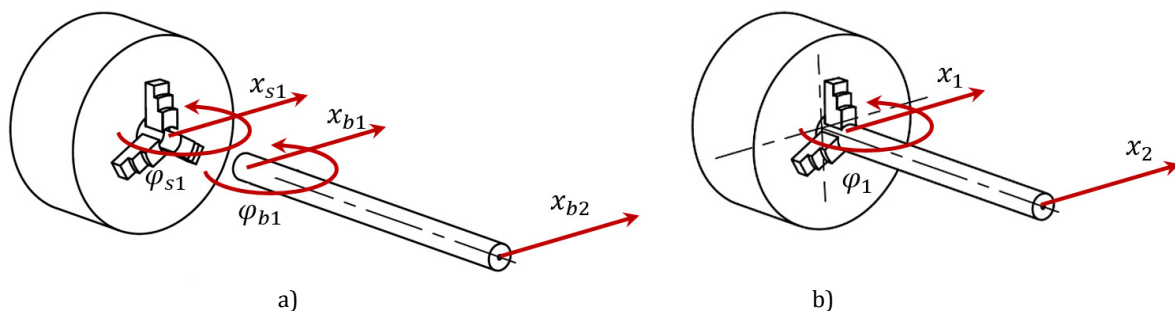


Fig. 2 a) The system subassemblies in the local coordinates; b) The coupled system

However, the use of receptance coupling for turning involves the difficulty of modelling the interaction between the workpiece and the three-jaw chuck. It is analogous to modelling machine (spindle and holder) tool joint for milling, for which the RCA method is most commonly used. Phenomena that occur when modelling such a fixture are presented, among others by Schmitz in [35]. To obtain correct calculation results, additional consideration of rotational degrees of freedom (RDOF) for the spindle - workpiece connection is required.

A Timoshenko beam model can be used in order to evaluate dynamic properties of the workpiece. First, information on geometry (diameter and length) and material properties are given, then boundary conditions are imposed (in this case “free - free”) and finally the required FRFs can be calculated.

Matrix equations describing the dynamic properties of system components for receptance coupling procedure are given as follows:

$$\begin{bmatrix} x_{s1} \\ \varphi_{s1} \end{bmatrix} = \begin{bmatrix} H_{s11} & H_{s12} \\ H_{s21} & H_{s22} \end{bmatrix} \cdot \begin{bmatrix} F_{s1} \\ M_{s1} \end{bmatrix} \tag{3}$$

$$\begin{bmatrix} x_{b1} \\ \varphi_{b1} \\ x_{b2} \\ x_{b3} \end{bmatrix} = \begin{bmatrix} H_{b11} & H_{b12} & H_{b13} & H_{b14} \\ H_{b21} & H_{b22} & H_{b23} & H_{b24} \\ H_{b31} & H_{b32} & H_{b33} & H_{b34} \\ H_{b41} & H_{b42} & H_{b43} & H_{b44} \end{bmatrix} \cdot \begin{bmatrix} F_{b1} \\ M_{b1} \\ F_{b2} \\ F_{b3} \end{bmatrix} \tag{4}$$

where: x - the x direction translational displacement, φ rotation angle, H_s - spindle FRFs, H_b - workpiece FRFs, F - force, M - torque.

For the spindle, the translational FRF H_{s11} can be determined by impulse testing. However, the RCA also requires the rotational FRFs (H_{s12} , H_{s21} , and H_{s22}) and the experimental determination of these is troublesome. The methods provided for the experimental determination of RDOFs in practice turn out to be sensitive to measurement inaccuracies or have numerous limitations resulting from the fact that they were mainly developed for tool-holder joints in milling. The issue of experimental determination of rotational FRFs for lathe applications was explored by Jasiewicz and Powalka in [33] where the Extended Inverse Receptance Coupling (EIRC) procedure was presented. The method assumes the determination of the FRF of a spindle using the inverse receptance coupling. The translational transfer functions of the system spindle - rod are evaluated experimentally (by impact testing) and the system is decoupled so that the FRF of the subassembly, i.e., the spindle, can be obtained. Moreover, in order to increase the calculation accuracy, the number of measurement configurations has been extended. The EIRC method was used to determine spindle FRFs.

While having the transfer functions of the workpiece and the spindle, the RCA can be performed to evaluate the machine tool system dynamic properties. The relationships between the points of the system constitute boundary conditions and equilibrium of forces. Moreover, local coordinates (for separate subassemblies) are replaced by the global ones (for a coupled system):

$$\begin{cases} x_{s1} = x_{b1} = x_1 \\ \varphi_{s1} = \varphi_{b1} = \varphi_1 \end{cases}, \quad \begin{cases} F_{s1} + F_{b1} = F_1 \\ M_{s1} + M_{b1} = M_1 \end{cases} \tag{5}$$

By including boundary conditions and equilibrium of forces in the spindle (Eq. 3) and beam (Eq.4) matrices the coupled system matrix equation is obtained:

$$\begin{bmatrix} x_1 \\ \varphi_1 \\ x_2 \\ x_3 \end{bmatrix} = T \cdot \begin{bmatrix} H_{b11} & H_{b12} & H_{b13} & H_{b14} \\ H_{b21} & H_{b22} & H_{b23} & H_{b24} \\ H_{b31} & H_{b32} & H_{b33} & H_{b34} \\ H_{b41} & H_{b42} & H_{b43} & H_{b44} \end{bmatrix} \cdot \begin{bmatrix} F_1 \\ M_1 \\ F_2 \\ F_3 \end{bmatrix} \tag{6}$$

In the matrix equation, Eq. 6, the dynamic properties of the coupled system are determined as the product of the matrix T and the matrix containing beam FRFs. The T matrix contains the spindle and beam FRFs and is given as:

$$T^{-1} = \begin{bmatrix} 1 + \frac{H_{b11}H_{S22} - H_{b12}H_{S12}}{H_{S11}H_{S22} - H_{S12}H_{S21}} & \frac{H_{b12}H_{S11} - H_{b11}H_{S12}}{H_{S11}H_{S22} - H_{S12}H_{S21}} & 0 & 0 \\ \frac{H_{b21}H_{S22} - H_{b22}H_{S12}}{H_{S11}H_{S22} - H_{S12}H_{S21}} & 1 + \frac{H_{b22}H_{S11} - H_{b21}H_{S12}}{H_{S11}H_{S22} - H_{S12}H_{S21}} & 0 & 0 \\ \frac{H_{b31}H_{S22} - H_{b32}H_{S12}}{H_{S11}H_{S22} - H_{S12}H_{S21}} & \frac{H_{b32}H_{S11} - H_{b31}H_{S12}}{H_{S11}H_{S22} - H_{S12}H_{S21}} & 1 & 0 \\ \frac{H_{b41}H_{S22} - H_{b42}H_{S12}}{H_{S11}H_{S22} - H_{S12}H_{S21}} & \frac{H_{b42}H_{S11} - H_{b41}H_{S12}}{H_{S11}H_{S22} - H_{S12}H_{S21}} & 0 & 1 \end{bmatrix} \quad (7)$$

Using the RCA procedure presented, FRFs of the machine tool system are obtained, which are the basis for predicting system stability and thus determining technological parameters to avoid vibrations during the machining.

2.3 Proposed machining parameters selection system

The beginning of the production process using a CNC machine tool is usually the analysis of technical documentation, i.e., technical drawing. Then, based on experience and knowledge about machining technology, the tools to perform the machining are selected. Regardless of the system in which it is developed, the machining program consists of commands describing the geometry of the part, allowing the tool path to be calculated. In addition, the operation of machine tool systems is controlled (e.g. spindle on/off, coolant activation, tool change), but also machining parameters must be selected. For turning, these include the feed, the cutting depth and the spindle rotational speed (or cutting speed). The manufacturers of the cutting inserts usually provide the recommended ranges to facilitate the selection of optimal parameters. However, this selection does not respect the dynamic properties of the machine tool-workpiece system and while machining of compliant parts, a small change in parameters can significantly change the cutting conditions i.e., affect whether the machining will be stable or whether vibration will occur. As an example, turning parts with a 30 mm diameter with a tool for which the manufacturer recommends the cutting speed from 335 to 450 m/min gives a spindle rotational speed in the range of 3350-4770 rpm. For a non-susceptible part (low length – diameter ratio), choosing a higher cutting speed will only mean higher machining efficiency with lower tool life, while for lower speeds, the cutting process will take longer, but more parts can be machined with the same tool insert. However, for the parts with high compliance, where the risk of chatter vibration occurrence during the cutting is particularly high, this issue becomes more complex. For such parts, it is justified to determine the stability lobes, representing the area of stable and unstable machining, as shown in Fig. 3.

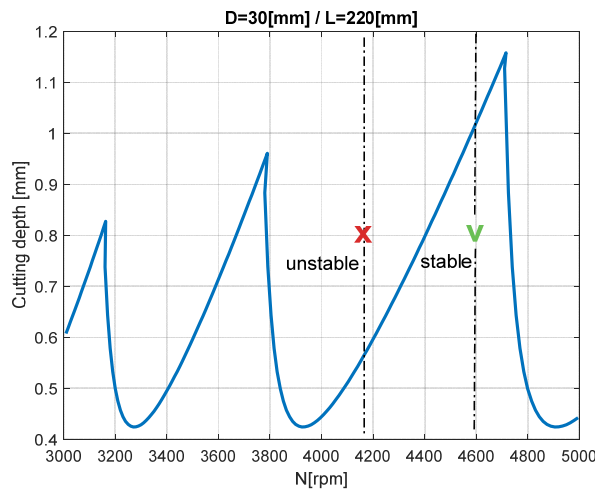


Fig. 3 Selection of the spindle rotational speed on stability lobes diagram

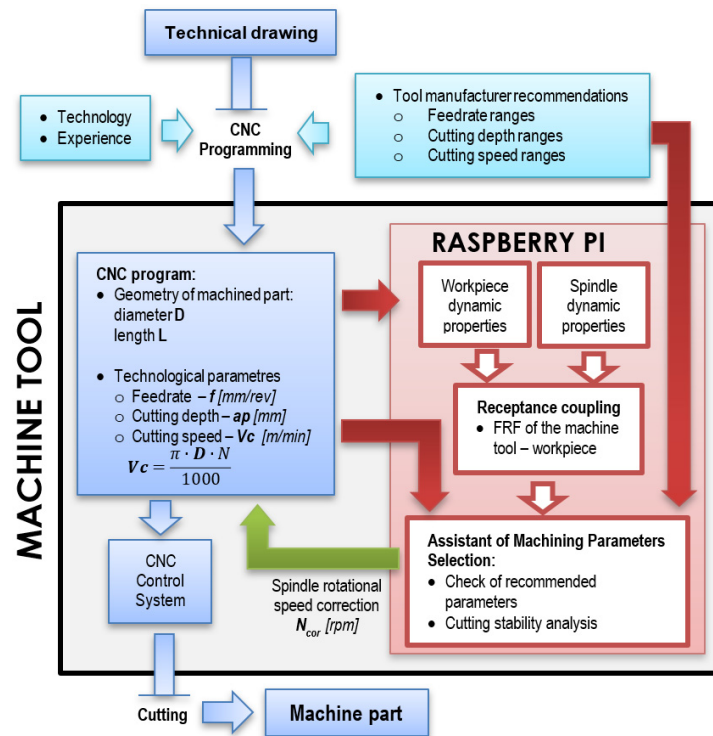


Fig. 4 The concept of machining parameters selection system

When analyzing the stability lobes, it can be seen that at a given cutting depth, for some spindle rotational speeds from the recommended speeds range, the machining will be stable, while for others the chatter vibration will occur. However, in practice, stability lobes are rarely determined before machining because they require an impact test to be carried out for a specific workpiece mounted in the spindle chuck. This can be done using the specialized measurement equipment, software and requires highly qualified staff. For these reasons, this solution is unprofitable for most small and medium-sized manufacturing companies. The proposed approach is a low-cost solution. Importantly, the system is integrated with the CNC system of the machine tool, which does not require performing any calculations on an external device. The concept of the system operation is presented in the algorithm in Fig. 4.

The machining program is developed in the same way as for standard machining, except that additionally, the cutting speed range must be specified. This may be the range provided by the tool insert manufacturer or resulting from the programmers experience. In order to determine the properties of the machine tool-workpiece system using RCA, it is necessary to have FRFs matrices of workpiece and spindle as the input. The spindle properties are determined using the EIRC method at the stage of system installation at the machine tool and the FRFs matrix is saved in the single-board computer (Raspberry Pi 4) memory. The dynamic properties of the workpiece are determined using the Timoshenko beam model with defined diameter and length. As presented earlier in section 2.1, the influence of the cutting force coefficient (related to material properties) is neglected, which in practice means that only "stable" rotational speeds of the spindle are sought. The cutting speed preselected in the machining program is compared with a set of "stable" spindle speeds and as a result, the system indicates the nearest spindle rotational speed that can contribute to chatter suppression.

In the presented form, the method is dedicated for turning and can be useful when machining slender workpieces. However, the concept of the machining parameters selection system can be also applied to milling operations. The receptance coupling method is commonly used to determine the dynamic properties of the spindle-holder-tool system. However, it should be emphasized that then the use of the system would be justified for machining with slender tools, which, due to their geometry, could be analytically modeled as a beam with a circular cross-section. For milling, there are also cases where thin-walled workpieces are machined. In these cases, the application of the presented system will not be helpful.

3. Results and discussion

3.1 Implementation of the system in the CNC control

The machining process assistant has been integrated with the FANUC CNC 31i model B control system. The FANUC control system is a part of a medium size lathe AFM TAE 35 "Hanka" manufactured by "Andrychowska Fabryka Maszyn DEFUM S.A". The speed selection algorithm has been implemented in a single-board computer Raspberry Pi 4B with Raspbian OS. Raspbian is an open-source operating system optimized for the SoC Raspberry Pi hardware. Raspbian includes a set of basic programs and utilities that allows the usage of the Raspberry Pi. MATLAB Coder was used to generating and deploying the C/C++ code on Raspberry Pi. Communication between RPi 4 and NC/PMC controller is implemented over the Ethernet. On the CNC control side, the Ethernet communication is implemented by the socket communication (TCP/IP) with FANUC Ethernet Board with the "Embedded Ethernet function". A diagram of the communication between the system components, with the libraries and software used, is shown in Fig. 5a.

The first step to integrating the developed algorithm into the CNC system was to develop an HMI user interface. The FANUC Picture software was used for this purpose. This software allows integrators to create customized operator screens and implement complex machining functions. Created screens are compiled and saved in the CNC Flash-ROM (FROM) memory. Properly prepared and uploaded compiled code is rendered by the CNC main processor on an HMI LCD screen. The prepared assistant main screen is shown in Fig. 5b.

After the machine is switched on, the start-up screen appears on the operator panel. The main screen allows the operator to switch directly to the assistant's settings or other machine functions. When the screen appears, the machining parameter fields are empty. The following parameters (necessary for the calculation) are required from the operator to use the system: workpiece length, workpiece diameter, cutting speed range, chosen cutting speed. Then entered parameters are saved in PMC memory in R (relay) data area. The next step necessary to integrate the proposed solution into the machine was to prepare the Raspberry Pi 4 single-board computer. First, using MATLAB Coder software, the MATLAB code of the developed algorithm for RPi 4 was generated. That code can be deployed and run standalone on the Raspberry Pi prepared by MATLAB Support Package for Raspberry Pi Hardware creator. The pseudo-code of the algorithm is given below.

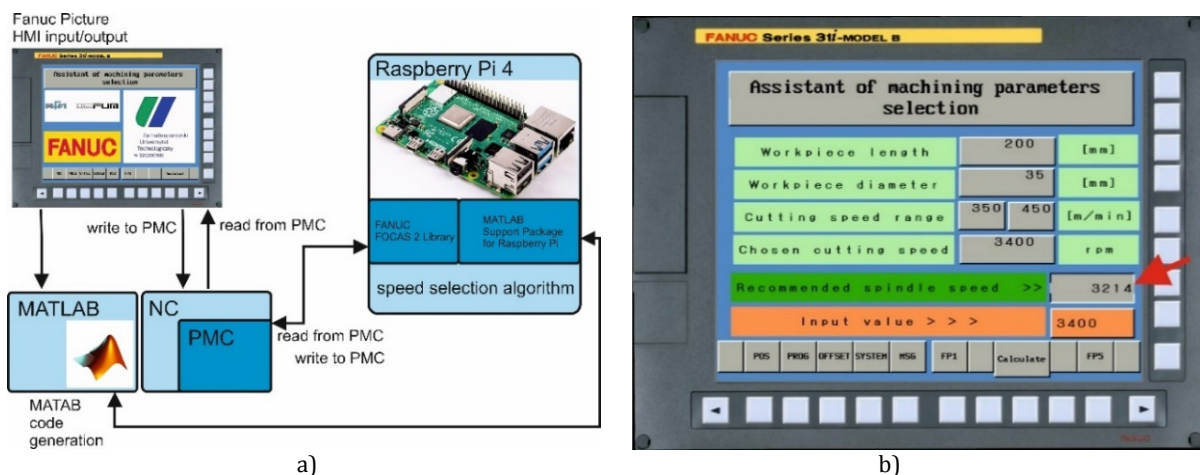


Fig. 5 a) Operating diagram of the integrated system for chatter suppression; b) Operator screen of the integrated system for chatter suppression

Table 1 Algorithm of spindle rotational speed selection pseudocode

Algorithm of spindle rotational speed selection	
	Input: $D, L, V_{c\ range}, V_{c\ select}$
	Output: N_{cor}
1	Read input data form CNC
2	Calculate beam FRFs matrix H_b
3	Load spindle FRFs matrix H_s
4	Synthesize FRFs of the coupled system using RCA
5	Determine the stability lobes and select stable speeds N_{stab}
6	Transform cutting speeds V_c to N rotational speeds
7	Find $N_{stab} \in N_{range}$ closest to N_{select} and assign as N_{cor}
8	Send N_{cor} to NC/PMC

D - workpiece diameter, L - workpiece length, $V_{c\ range}/N_{range}$ - cutting speed/spindle rotational speed range, $V_{c\ select}/N_{select}$ - cutting speed/spindle rotational speed selected in machining program, N_{stab} - stable spindle rotational speeds, N_{cor} - recommended spindle rotational speed

FANUC FOCAS2 for ARM libraries were used to read/write values of variables from/to R registers of the PMC controller. The use of these libraries allows creating custom programs and cycles that are not available by default in the CNC system. FOCAS2 libraries provide direct access to the variables, parameters, registers in NC/PCM memory areas, control of the machine axis, perform complex calculations and connecting external devices. Establishing communication between the CNC controller and SoC requires creating the library handle and the TCP/IP connection by passing a CNC's IP address parameter to the `cnc_allclibhdl3` function. Acquired library handle number must be known until the application program terminates. The actual instance handle number is required as an argument in every call CNC/PMC Data window library function.

The following functions were used to read and write values using RPi 4 via Ethernet connection: `cnc_rdunsolicprm2` (to read parameters), `cnc_wrunsolicprm2` (to write parameters). It is necessary to execute `cnc_wrunsolicprm2` and make the parameter effective before reading the parameters. The outline of the processing sequence of a FOCAS2 application is shown on Fig. 6a. In the prototype system phase it was necessary to use the NCGuide CNC simulator. FANUC NCGuide is a software for simulating real CNC machine (G-code and PLC/PMC code). It allows for real programming, operation, and maintenance environment without using a production machine tool. For a better understanding overview of the complete control system architecture with connections is presented in Fig. 6b.

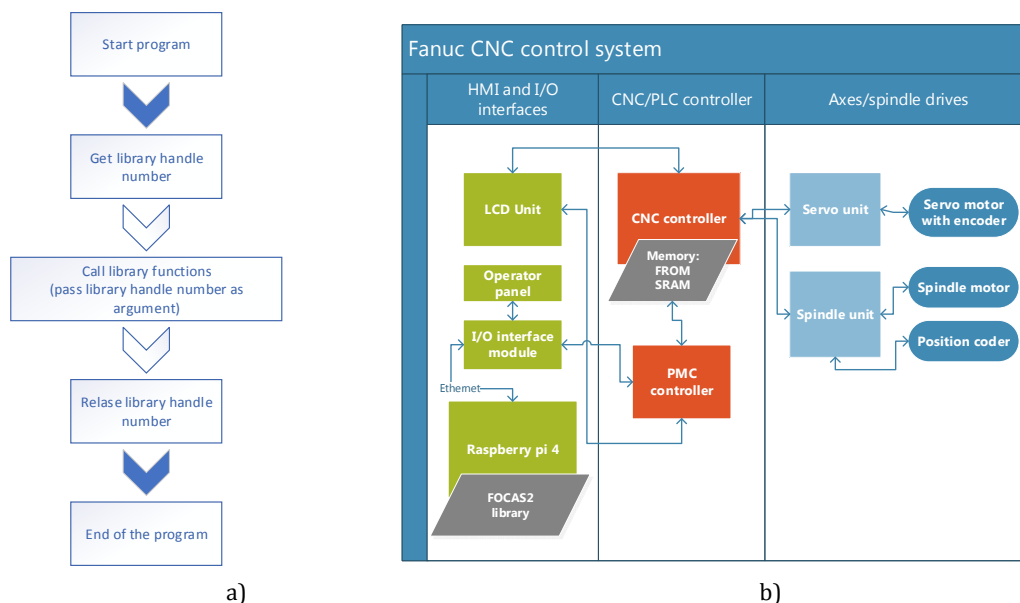


Fig. 6 a) FOCAS2 application processing sequence; b) Structure of Fanuc CNC control system

At the prototype stage of the solution, all necessary parameters must be entered to calculate the recommended settings. In the next version of the system, the number of parameters required to enter will be reduced. A special algorithm will be prepared which will download cutting speed and the diameter of the workpiece information from G-code. The main advantage of the developed solution is that it: does not change the existing system functionality, increases user convenience by limiting the amount of data entered, additional equipment allowing to perform calculations is fully integrated with the CNC system.

3.2 Experimental tests

In this subsection, the experimental test results were presented to validate the proposed solution. The system was integrated into a AFM TAE35 CNC lathe on which cutting tests on the slender workpiece were carried out. The workpiece was a circular cross section rod of a diameter of 40 mm, length 250 mm and the material was a machining steel 1.0715 (11SMn30). The SVJCL 2020-16 cutting tool was used, equipped with VCMT 160402-SM cutting insert.

The initial step of the procedure was to determine the dynamic properties of the machine tool-workpiece system using the RCA. For the analysis of machining stability, the most significant is the first vibration mode, which for the selected system was at a frequency of 223 Hz. Then, the stability lobes diagram for the selected system was evaluated, as given in Fig. 7. The cutting test plan assumed the machining of two 10mm long sections at the end of the workpiece. Due to the lack of support with the tailstock, the rigidity of the system is the lowest at the end, and therefore the risk of chatter vibrations is the greatest. In order to set the same cutting depth for both sections, without passing through the machined surface, "Section 1" (closer to the face) has been pre-machined. The effectiveness of the proposed procedure consisted of machining Section 1, with the spindle rotational speed set arbitrarily, whereas for "Section 2" the spindle speed has been corrected using the presented machining parameters selection system. For both sections, the cutting depth and the feed value remained unchanged. Fig. 8 presents the results of one of the performed experiments.

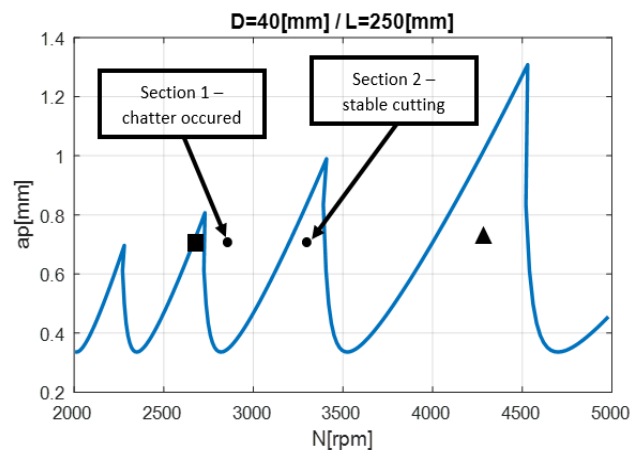


Fig. 7 The Stability lobes for the selected machine tool - workpiece system

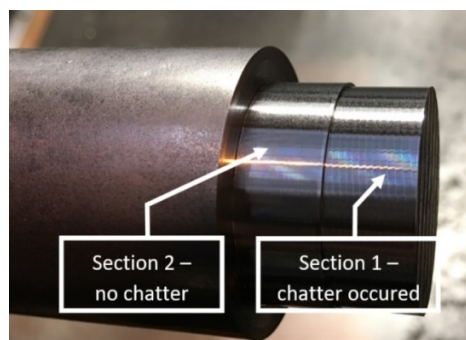


Fig. 8 Experimental test result - machined surfaces

For the “Section 1”, the selected cutting speed was 290 m/min, which for the 32 mm diameter gives a rotational speed of approx. 2800 rpm. The cutting depth a_p was 0.7 mm and feed per revolution 0.05 mm. The analysis of the stability lobes diagram in Fig. 9 gives that for the machine tool-workpiece system, the selection of this rotational spindle speed may lead to chatter vibration occurrence during the machining. However, this cannot be stated without the stability lobes for the workpiece setting, normally not determined in industrial applications. Thus, despite the selection of the cutting speed from the range given by the cutting insert producer, the chatter marks appeared on the machined surface in “Section 1”. The cutting in “Section 2” was performed with the spindle rotational speed proposed by the machining parameters selection system. Setting the same feedrate and cutting depth as for “Section 1”, but correction of the spindle rotational speed to 3250 rpm, allowed to obtain a stable, chatter-free machining process. As can be seen in Fig. 7, in the presented rotational speed range, other stable speeds can be selected for determined cutting depth. However, by choosing a speed of 4200 rpm (pointed as triangle), the cutting speed is outside the speed range proposed by the tool manufacturer, which may result in accelerated tool wear or even cutting-edge damage. For a lower speed, e.g. 2600 rpm (pointed as a square), the machining performance is lower, and there is an increased risk of going outside the stable machining area. At the surface machined with the corrected spindle rotational speed no chatter marks were observed, which indicates that the manufacturing process could be successfully continued.

4. Conclusion

The paper presents the issue of selecting the cutting parameters for compliant workpieces turning. The selection of these parameters is practically based on ranges provided by the cutting tools producers or on the experience of the programmer. In most cases, this approach is sufficient and only the manufacturing efficiency and tool life depend on the choice made. However, the selection of rotational speed and cutting depth is particularly important for compliant parts machining, as it can affect the occurrence of unwanted chatter vibrations. The algorithm implemented on a single board computer integrated into the CNC control system presented in the article allows supporting the task of selecting these parameters. The determination of machining stability is based on the analysis of the frequency response function of the workpiece mounted in the lathe. A significant advantage of the presented system compared to other research works on turning stability is the omission of impulse tests by application of the receptance coupling approach (most commonly used for milling). The dynamic properties of the machine tool-workpiece system determined by this method are a substantial part of the standard stability lobes determination procedure. However, in order to simplify the procedure for selecting technological parameters, the influence of the cutting force factor was omitted. As a result, it only allows searching for the “stable rotational speeds” without setting a cutting depth limit that may be considered a drawback of the system. Nevertheless, in many cases, this simplification proves to be sufficient and may be useful in industrial practice while selecting the machining parameters. An important advantage of the proposed solution is integration with the CNC control system. The system is operated from the operator panel and only calculations are carried out on an external single board computer. This solution does not require an independent computer station outside the machine tool and therefore is a low-cost solution. Moreover, what is particularly important from a practical point of view, it can be installed both on brand new and already used machine tools already working in the industry. Due to the use of a single-board computer Raspberry Pi, the system is not a closed solution and new options and improvements can easily be introduced. Although the system presented in the article refers to turning, in some applications, it can also be adapted to milling. The following studies will concern, among others, extending the stability analysis with cutting depth and introducing the system interface improvements.

Acknowledgement

Research carried out on research apparatus purchased as part of the project No. RPZP.01.03.00-32-0004/17. Project co-financed by the European Union from the European Regional Development Fund under the Regional Operational Program of the West Pomeranian Voivodeship 2014-2020. Project co-financed by the Ministry of Science and Higher Education

References

- [1] Tlustý, J., Polacek, M. (1963). The stability of machine tools against self-excited vibrations in machining, In: *Proceedings of the ASME International Research in Production Engineering Conference*, Pittsburg, USA, 465-474.
- [2] Tobias, S.A., Fishwick, W. (1958). Theory of regenerative machine tool chatter, *The Engineer*, Vol. 205, No. 7, 199-203.
- [3] Altintas, Y. (2012). *Manufacturing automation: Metal cutting mechanics, machine tool vibrations, and CNC design*, Cambridge University Press, New York, USA, doi: [10.1017/CBO9780511843723](https://doi.org/10.1017/CBO9780511843723).
- [4] Powalka, B., Jemielniak, K. (2015). Stability analysis in milling of flexible parts based on operational modal analysis, *CIRP Journal of Manufacturing Science and Technology*, Vol. 9, 125-135, doi: [10.1016/j.cirpj.2014.11.003](https://doi.org/10.1016/j.cirpj.2014.11.003).
- [5] Kalinski, K.J., Galewski, M.A. (2011). Chatter vibration surveillance by the optimal-linear spindle speed control, *Mechanical Systems and Signal Processing*, Vol. 25, No. 1, 383-399, doi: [10.1016/j.ymssp.2010.09.005](https://doi.org/10.1016/j.ymssp.2010.09.005).
- [6] Siddhpura, M., Paurobally, R. (2012). A review of chatter vibration research in turning, *International Journal of Machine Tools and Manufacture*, Vol. 61, 27-47, doi: [10.1016/j.ijmactools.2012.05.007](https://doi.org/10.1016/j.ijmactools.2012.05.007).
- [7] Munoa, J., Beudaert, X., Dombovari, Z., Altintas, Y., Budak, E., Brecher, C., Stepan, G. (2016). Chatter suppression techniques in metal cutting, *CIRP Annals*, Vol. 65, No. 2, 785-808, doi: [10.1016/j.cirp.2016.06.004](https://doi.org/10.1016/j.cirp.2016.06.004).
- [8] Urbikain, G., Olvera, D., López de Lacalle, L.N., Beranoaguirre, A., Elías-Zúñiga, A. (2019). Prediction methods and experimental techniques for chatter avoidance in turning systems: A review, *Applied Sciences*, Vol. 9, No. 21, Article No. 4718, doi: [10.3390/app9214718](https://doi.org/10.3390/app9214718).
- [9] Dunaj, P., Berczyński, S., Dolata, M. (2019). Modelling machine tool rocking vibrations using reduced order models, In: Rusiński, E., Pietrusiak, D. (eds.), *Proceedings of the 14th International Scientific Conference: Computer Aided Engineering, CAE 2018, Lecture Notes in Mechanical Engineering*, Springer, Cham, Switzerland, 183-190, doi: [10.1007/978-3-030-04975-1_22](https://doi.org/10.1007/978-3-030-04975-1_22).
- [10] Dunaj, P., Okulik, T., Powalka, B., Berczyński, S., Chodźko, M. (2019). Experimental investigations of steel welded machine tool bodies filled with composite material, In: Gapiński, B., Szostak, M., Ivanov, V. (eds.), *Advances in Manufacturing II, MANUFACTURING 2019. Lecture Notes in Mechanical Engineering*, Springer, Cham, Switzerland, 61-69, doi: [10.1007/978-3-030-16943-5_6](https://doi.org/10.1007/978-3-030-16943-5_6).
- [11] Dunaj, P., Berczyński, S., Chodźko, M. (2020). Method of modeling steel-polymer concrete frames for machine tools, *Composite Structures*, Vol. 242, Article No. 112197, doi: [10.1016/j.compstruct.2020.112197](https://doi.org/10.1016/j.compstruct.2020.112197).
- [12] Powalka, B., Okulik, T. (2012). Dynamics of the guideway system founded on casting compound, *The International Journal of Advanced Manufacturing Technology*, Vol. 59, 1-7, doi: [10.1007/s00170-011-3472-6](https://doi.org/10.1007/s00170-011-3472-6).
- [13] Dunaj, P., Berczyński, S., Miądlicki, K., Irska, I., Niesterowicz, B. (2020). Increasing damping of thin-walled structures using additively manufactured vibration eliminators, *Materials*, Vol. 13, No. 9, Article No. 2125, doi: [10.3390/ma13092125](https://doi.org/10.3390/ma13092125).
- [14] Parus, A., Powalka, B., Marchelek, K., Domek, S., Hoffmann, M. (2013). Active vibration control in milling flexible workpieces, *Journal of Vibration and Control*, Vol. 19, No. 7, 1103-1120, doi: [10.1177/1077546312442097](https://doi.org/10.1177/1077546312442097).
- [15] Brecher, C., Manoharan, D., Ladra, U., Köpken, H.-G. (2010). Chatter suppression with an active workpiece holder, *Production Engineering – Research and Development*, Vol. 4, 239-245, doi: [10.1007/s11740-009-0204-y](https://doi.org/10.1007/s11740-009-0204-y).
- [16] Altintas, Y., Budak, E. (1995). Analytical prediction of stability lobes in milling, *CIRP Annals*, Vol. 44, No. 1, 357-362, doi: [10.1016/S0007-8506\(07\)62342-7](https://doi.org/10.1016/S0007-8506(07)62342-7).
- [17] Altintas, Y., Weck, M. (2004). Chatter stability of metal cutting and grinding, *CIRP Annals*, Vol. 53, No. 2, 619-642, doi: [10.1016/S0007-8506\(07\)60032-8](https://doi.org/10.1016/S0007-8506(07)60032-8).
- [18] Ahmadi, K., Ismail, F. (2011). Analytical stability lobes including nonlinear process damping effect on machining chatter, *International Journal of Machine Tools and Manufacture*, Vol. 51, No. 4, 296-308, doi: [10.1016/j.ijmactools.2010.12.008](https://doi.org/10.1016/j.ijmactools.2010.12.008).
- [19] Schmitz, T.L., Smith, K.S. (2019). *Machining dynamics: Frequency response to improved productivity*, Springer, New York, USA, doi: [10.1007/978-3-319-93707-6](https://doi.org/10.1007/978-3-319-93707-6).
- [20] Mahdavinejad, R. (2005). Finite element analysis of machine and workpiece instability in turning, *International Journal of Machine Tools and Manufacture*, Vol. 45, No. 7-8, 753-760, doi: [10.1016/j.ijmactools.2004.11.017](https://doi.org/10.1016/j.ijmactools.2004.11.017).
- [21] Dunaj, P., Marchelek, K., Chodźko, M. (2019). Application of the finite element method in the milling process stability diagnosis, *Journal of Theoretical and Applied Mechanics*, Vol. 57, No. 2, 353-367, doi: [10.15632/jtam-pl/104589](https://doi.org/10.15632/jtam-pl/104589).
- [22] Dunaj, P., Berczyński, S., Chodźko, M., Niesterowicz, B. (2020). Finite element modeling of the dynamic properties of composite steel-polymer concrete beams, *Materials*, Vol. 13, No. 7, Article No. 1630, doi: [10.3390/ma13071630](https://doi.org/10.3390/ma13071630).
- [23] Jasiewicz, M., Powalka, B. (2016). Receptance coupling for turning with a follower rest, In: Kleiber, M., Burczynski, T., Wilde, K., Gorski, J., Winkelmann, K., Smakosz, L. (eds.), *Advances in Mechanics: Theoretical, Computational and Interdisciplinary Issues*, CRC Press Taylor & Francis, London, United Kingdom, 245-248, doi: [10.1201/b20057](https://doi.org/10.1201/b20057).

- [24] Park, S.S., Altintas, Y., Movahhedy, M. (2003). Receptance coupling for end mills, *International Journal of Machine Tools and Manufacture*, Vol. 43, No. 9, 889-896, doi: [10.1016/S0890-6955\(03\)00088-9](https://doi.org/10.1016/S0890-6955(03)00088-9).
- [25] Jasiewicz, M., Powalka, B. (2018). Prediction of turning stability using receptance coupling, In: *Proceedings of the 22nd International Conference on Computer Methods in Mechanics, AIP Conference Proceedings*, Vol. 1922, Lublin, Poland, doi: [10.1063/1.5019090](https://doi.org/10.1063/1.5019090).
- [26] Mia, M., Królczyk, G., Maruda, R., Wojciechowski, S. (2019). Intelligent optimization of hard-turning parameters using evolutionary algorithms for smart manufacturing, *Materials*, Vol. 12, No. 6, Article No. 879, doi: [10.3390/ma12060879](https://doi.org/10.3390/ma12060879).
- [27] Chiu, H.-W., Lee, C.-H. (2020). Intelligent machining system based on CNC controller parameter selection and optimization, *IEEE Access*, Vol. 8, 51062-51070, doi: [10.1109/ACCESS.2020.2980286](https://doi.org/10.1109/ACCESS.2020.2980286).
- [28] Jasiewicz, M., Miądlicki, K., Powalka, B. (2018). Assistance of machining parameters selection for slender tools in CNC control, In: *Proceedings of Mechatronics systems and materials 2018, AIP Conference Proceedings*, Vol. 2029, Zakopane, Poland, Article No. 020024, doi: [10.1063/1.5066486](https://doi.org/10.1063/1.5066486).
- [29] Sun, Y., Liu, C., Sun, L., Xiong, Z., Zhu, X. (2020). Chatter detection with beat effect based on beat frequency estimation, *IEEE Transactions on Automation Science and Engineering*, Vol. 17, No. 4, 1-9, doi: [10.1109/TASE.2020.3001623](https://doi.org/10.1109/TASE.2020.3001623).
- [30] Liu, X., Su, C.-Y., Li, Z., Yang, F. (2018). Adaptive neural-network-based active control of regenerative chatter in micromilling, *IEEE Transactions on Automation Science and Engineering*, Vol. 15, No. 2, 628-640, doi: [10.1109/TASE.2017.2667709](https://doi.org/10.1109/TASE.2017.2667709).
- [31] Ma, H., Guo, J., Wu, J., Xiong, Z., Lee, K.-M. (2020). An active control method for chatter suppression in thin plate turning, *IEEE Transactions on Industrial Informatics*, Vol. 16, No. 3, 1742-1753, doi: [10.1109/TII.2019.2924829](https://doi.org/10.1109/TII.2019.2924829).
- [32] Chen, F., Hanifzadegan, M., Altintas, Y., Lu, X. (2015). Active damping of boring bar vibration with a magnetic actuator, *IEEE/ASME Transactions on Mechatronics*, Vol. 20, No. 6, 2783-2794, doi: [10.1109/TMECH.2015.2393364](https://doi.org/10.1109/TMECH.2015.2393364).
- [33] Tomac, N., Sørby, K., Doboviček, S. (2018). Formation of built-up layer on the tool in turning operation of magnesium alloys, *Tehnički Vjesnik – Technical Gazette*, Vol. 25, No. 3, 940-943, doi: [10.17559/TV-20151027130036](https://doi.org/10.17559/TV-20151027130036).
- [34] Świć, A., Wołos, D., Gola, A., Smidova, N. (2019). Accuracy control in the process of low-rigidity elastic deformable shafts turning, *Tehnički Vjesnik – Technical Gazette*, Vol. 26, No. 4, 927-934, doi: [10.17559/TV-20170924230143](https://doi.org/10.17559/TV-20170924230143).
- [35] Schmitz, T. (2020). Modal interactions for spindle, holders, and tools, *Procedia Manufacturing*, Vol. 48, 457-465, doi: [10.1016/j.promfg.2020.05.069](https://doi.org/10.1016/j.promfg.2020.05.069).

Simulation-based time evaluation of basic manual assembly tasks

Turk, M.^a, Pipan, M.^{a,*}, Šimic, M.^a, Herakovič, N.^a

^aUniversity of Ljubljana, Faculty of Mechanical Engineering, Slovenia

ABSTRACT

The paper presents a simple simulation model of the lifting procedure that can be used to predict the total time required for the sequence of basic manual assembly tasks depending on the various parameters of the load and with regard to the workers' health. The aim of the research is to determine the appropriateness of using simulation tool for (re)setting time standards for manual assembly tasks. An avatar in the simulation model performs sequences of tasks with a handling mass of up to 20.5 kg. The individual times obtained from the simulation model were analysed and compared with several time prediction methods and validated in laboratory environment. An analysis of the influence of different load parameters on the total time was also performed. Dependency is mostly linear, so from the practitioner point of view, we can predict with reasonable certainty the total time for any sequence of manual assembly tasks for every size and mass of the box. Based on the results we can confirm that simulation tool JACK is suitable not only for ergonomic analyses but also for setting time standards for the workers. Furthermore, with the simulation tool we analyse the process and get the accurate results in shorter time compared to other mentioned methods.

© 2020 CPE, University of Maribor. All rights reserved.

ARTICLE INFO

Keywords:
Assembly;
Manual task;
Work-job design;
Time analysis;
Jack simulation;
Avatar

**Corresponding author:*
miha.pipan@fs.uni-lj.si
(Pipan, M.)

Article history:
Received 20 April 2020
Revised 25 September 2020
Accepted 29 September 2020

1. Introduction

Many tasks at industrial assembly workplaces still require manual work that includes a variety of activities such as loading and unloading, pushing and pulling, and carrying tasks that require manual handling of goods and materials (MMH) [1, 2]. When designing jobs and products the aggregated information on processes, tools, machines, subjects of work, tasks and operators must be taken into account, limitations, which are often conflicting, must be met and a design must be generated, which will be acceptable for all parties involved [3]. In order to address workplace design from an ergonomic and health point of view, it is necessary to predict the times required for a worker to complete individual work tasks. These times are important for determining expected productivity, planning staff and material requirements in the workplace, conducting ergonomic assessments, reducing work-related musculoskeletal disorders (WMSDs), etc. [4-7]. They are usually predicted by the use of Predetermined Motion Time Systems (PMTS), such as the Methods Time Measurement (MTM) and the Maynard Operation Sequence Technique (MOST) [8, 9]. Digital Human Models (DHMs) are effective design tools for visualizations, time analyses and ergonomic evaluation of user and workplace interactions in terms of reach, clearance, visibility and comfort. There is a gap between the use of DHM tools for ergonomic assessment of workplaces and the use of DHM tools for advanced time analysis according to dif-

ferent load parameters and avatar poses. To close this gap, there are still challenges to be addressed [7, 10], and our research is focused in this direction.

In this paper we describe step by step a time assessment methodology for basic assembly tasks in a sequence (multitasking, combined tasks) used in an industrial environment. We compare time reports obtained in the conventional way with the MTM method, the simulation tool Siemens JACK 9.0, a laboratory experiment and a new biomechanical time prediction model developed and presented by Harrari *et al.* [8]. In our case study we focused on whether the simulation done in Siemens Jack produces the same results as the MTM method or whether there are parameters (trajectories of hands, banding routine, dimension and mass of the lifting object) that lead to different times in Siemens Jack simulation compared to the MTM method. In addition, the results of the time analysis were also calculated with the new method NTPM developed by Harrari *et al.* [8] and verified by a laboratory experiment. Other research has also been carried out in the field of time and ergonomics analysis and DHM simulations of the lifting procedure, both in the studies of single tasks and as well in combined tasks.

Firstly, in the literature review, we first concentrate on individual tasks of the lifting procedure and analyses that other researchers have dealt with, taking into account different parameters, gender, mass of the load, etc. Secondly, we will discuss studies on combined tasks. We will also review the case studies of DHM simulation and finally focus on the studies that are most relevant to our case.

Padula *et al.* [11] studied the DHM simulation of trunk movement when lifting the load to different heights. The experiment was performed on different population groups (female, male, students, and workers with and without musculoskeletal symptoms). Martinez *et al.* focused on the study of gender differences in upper limb technique during a lifting task of a 6 or 12 kg box from hip to eye level [12]. Other researchers focused on studying the influence of box weight or handling height on the biomechanical exposure of workers [1, 8, 9], while others investigated the correlation between manual handling and injuries [13-16], and some of them focused on the maximum acceptable weight of a lift (MAWL) and lifting frequency [17, 18]. At this point, it must first be emphasized that our case is focused on lifting and lowering tasks, which are only part of the MMH tasks, and that we have not concentrated on a part of the MMH that includes, for example, pushing and pulling tasks. Secondly, our study is based on the combination of the basic tasks of MMH, as we named it as sequence of tasks or lifting procedure. In the area of combined tasks it is necessary to mention the studies for Straker *et al.* [19]. They combined basic manual handling activities such as pulling, lifting, carrying, lowering and pushing and investigated how the risk of such combined tasks could be assessed. The aim of the study was to compare the risks assessed in single manual handling tasks with the risks of combination tasks according to Maximum Acceptable Weights (MAWs). They concluded that the risk assessment of combined manual handling tasks using MAW measures cannot be performed accurately when using the risk assessments of isolated single tasks. In [20] they focused on ratings of discomfort, exertion and heart rate and concluded that combination task discomfort Sum, Rating of Perceived Exertion and heart rate measures were different to measures of the component of single tasks.

Different DHM tools are used to speed up a manual workplace and to use "what-if" scenarios for time and ergonomic analyses. Many researchers have investigated DHM tool in different situations and industrial fields. Several studies have reported on the use of DHMs in the automotive, aerospace and other industries [2, 10, 21]. There is some research that uses DHM tools as safety training methods [7]. There are also reports of disadvantages in the use of DHM software for example when working with the workers with disabilities [22]. The study focuses mainly on working environments where manual work is presented with the aim of creating new classifications of disabilities related to a manufacturing environment. The problem is that the well-known ergonomic software packages of DHM do not include workers with disabilities for their ergonomic and time analysis. Other studies, more related to our case study, have focused on work cycle time and time prediction models [6, 8, 9, 23]. In many cases, the PMTS method does not accurately take into account the physiological and biomechanical aspects in time predictions. Especially in cases of lifting, carrying and lowering objects, the PMTS method predicts a shorter time period than a worker is capable of performing without health consequences in the future.

The most relevant studies regarding our research on time prediction models can be found in [8, 9, 24]. The [9] covers the MTM experiment and MOST analysis, but compared to this research it does not include computer simulations. The focus is on the whole body of the worker and the aim was to develop a new time prediction model. The [8] deals with the design of a workplace with manual material handling tasks. It considers both productivity and ergonomics. It includes the DHM simulation tool and proposes a new time prediction model to be used in our study. The same new time prediction model was used in [24], as a parameter for optimization of the productivity. In [24] authors presents an innovative framework for formulating workplace design as an optimization problem that maximizes productivity while maintaining ergonomic assessment values below commonly used thresholds.

Based on the literature review described above, we can conclude that the DHM simulation tool is mainly used as an ergonomic assessment tool, rather than as a tool that can be used in the process of workplace design based on time analysis and different load's parameters. Therefore, the main idea of our research is the usage of the DHM for both, ergonomic and time-based analysis, which can be performed in much shorter time compared to other mentioned approaches and with equal time prediction accuracy, is useful and reliable method.

2. Materials, methods, and experimental work

In this section the methodology, the methods of study and the analysis of our research are presented step by step. The research is divided into two parts, as shown in Fig. 1. The first part is a comparison of the total times obtained with four methods (MTM, Jack tool, NTPM, laboratory experiment) and a second part is a study of the influential parameters. Besides the overview, the methodology is divided into three sub-sections, the first part is a case study presenting the basic assembly operations. The next subsection is a comparison of the methods and an explanation why and how different basic operations can be compared with each other using different methods. The third part deals with the influential parameters of the box.

2.1 Overview of the research approach

The overview of the research approach and the steps of the case study are shown in Fig. 1. The study is divided into two parts. The first part includes a time analysis of the total time of the lifting procedure. The lifting object is the cube-shaped box with evenly distributed weight. The dimension (height/width/depth) of the box is 400 mm and a mass is 13.5 kg. The lifting height is 800 mm. The results of the time analysis were obtained with four different methods (Jack simulation, MTM method, NTPM [8], laboratory experiment). The study is only a simulation study, which aims to test the difference in the results due to the different time prediction models. The results were compared to determine if simulation is a suitable tool for designing work tasks. We know that the simulation tool enables "what-if" scenarios, which only facilitates and speeds up the planning of work tasks, but if the tools were also suitable for design, it would make it easier to set time standards for workers. The simulation was carried out in the Siemens Jack program based on the MTM method. For comparison, we calculated time standards for the sequence of tasks with the classic MTM method and the NTPM method, which extends certain times for the execution of tasks by the worker with the aim of not causing injuries or WMSDs. The results were also verified by a laboratory experiment with 10 healthy students. All subjects were recruited on a voluntary basis.

The total time of lifting procedure TT [s] (manual material handling process) consists of walking, banding, applying force, lifting, carrying (walking with box), putting the box (lowering) and posing in neutral position.

$$TT = t_{walk} + t_{band} + t_{apply} + t_{lift} + t_{carry} + t_{lower} + t_{pose} \quad (1)$$

where t_{walk} , t_{band} , t_{apply} , t_{lift} , t_{carry} , t_{lower} , and t_{pose} , are the times required to walk, to band for the mass (box) reach it and grasp it, to apply force to the mass, to lift the mass, to carry the mass (walking with mass), to lower the mass (put and release the mass) and to pose the body in neutral position.

Individual times has different dependency among parameters. Following equations (Eq. 1 – Eq. 8) show functional dependency for each time, which participate in Eq. 1.

$$t_{walk} = f\{d_{walking}, d_{turning}, v\} \tag{2}$$

where $d_{walking}$ [mm] is distance walk without object, $d_{turning}$ [mm] is distance when turning without object and v [m/s] is velocity of walking/turning. Walking and turning distances are exclusive, so we use $d_{walking}$ for task walk and $d_{turning}$ for task turn the body.

$$t_{band} = f\{\Delta\theta_{TB}, \omega_{rB}, d_R; dim\} \tag{3}$$

where $\Delta\theta_{TB}$ [°] is trunk extension angle, ω_{rB} is angular velocity [s⁻¹] of body, d_R is reach distance [mm] towards the object, and dim [mm] is dimension of the box, which is important for grasp operation.

$$t_{apply} = f\{m, dim, \mu\} \tag{4}$$

where m [kg] is weight of the object, dim [mm] is dimension of the object and μ is friction coefficient between two materials (skin, cardboard).

$$t_{lift} = f\{m, \Delta\theta_T, \omega_r(m)\} \tag{5}$$

where m [kg] is weight of the object, $\Delta\theta_{TB}$ [°] is trunk extension angle and ω_r is angular velocity [s⁻¹] of body and object weight.

$$t_{carry} = f\{m, d, v\} \tag{6}$$

where m [kg] is weight of the object, d [mm] is distance of carrying (walking with object), and v [m/s] is velocity of walking.

$$t_{lower} = f\{m, \Delta\theta_T, \omega_r(m), pos\} \tag{7}$$

where m [kg] is weight of the object, $\Delta\theta_{TB}$ [°] is trunk extension angle and ω_r is angular velocity [s⁻¹]. “*pos*” is position of the object when releasing it. Position contains class of fit, case of symmetry, ease of handling.

$$t_{pose} = f\{joints\ location\} \tag{8}$$

where “joints locations” means the differences between current joint location and joint location of neutral pose.

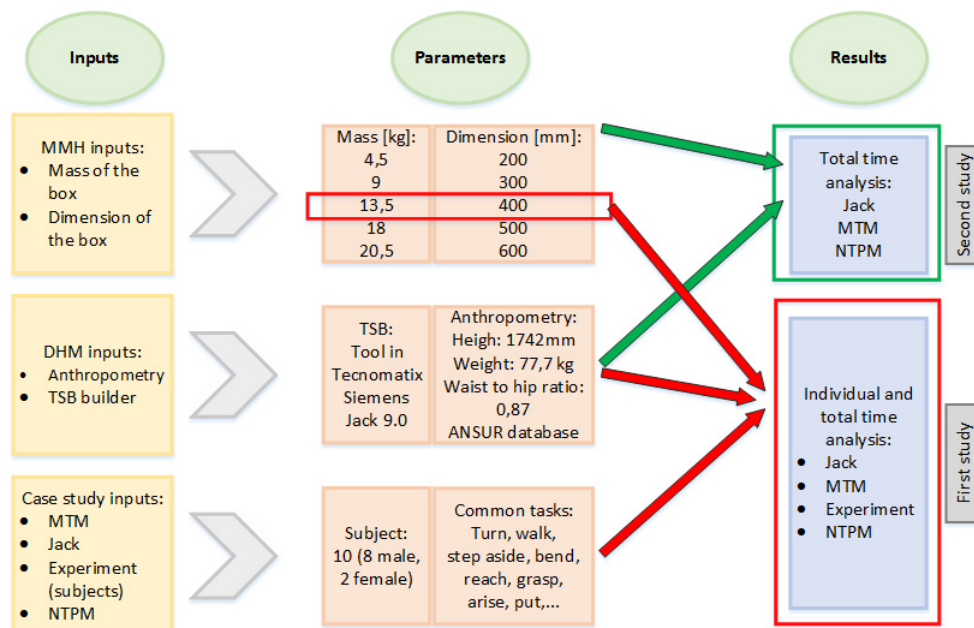


Fig. 1 Overview of the inputs, the parameters and the results

The second part of the study includes the total time analysis with different parameters of the load during the lifting and lowering procedure. The parameters we varied are the dimension (height, width, depth) and the mass of the load (box). The time analysis was performed by simulation in the Jack program, by the conventional MTM method and by the NTPM method. The lifting height is 800 mm. The analysis covers five different dimensions of boxes, i.e. cubes without handles. The dimensions of the cubes (*height = width = depth*) are 200 mm, 300 mm, 400 mm, 500 mm, 600 mm. The weight of the box is evenly distributed. Mass of the box covers the following values: 4.5 kg, 9 kg, 13.5 kg, 18 kg and 20.5 kg. The comparison was therefore made with 25 different combinations of parameters.

The purpose of the first study on the review of total working time using four methods and tools is to determine whether the simulation tool is suitable for work design and the definition of the relevant standards/norms for workers in industry. This would significantly reduce the time needed to design a workplace and make workplaces more worker-friendly. The purpose of determining the relationship between box dimensions, box weight and time according to the MTM, NTPM and Jack methods is to test whether the simulation tool shortens or lengthens the overall working time when the avatar lifts or lowers a larger or heavier box (Study 2).

2.2 Case study

Our case study is an example of a simple and common problem for the manual assembly process in industry (warehouses, manual assembly area, logistics, etc.). The definition of the problem is a lifting procedure and its time analyses. Most manual assembly process in industry, especially order-picking systems used in practice, are manual “picker to part” systems, and more than 80 % of all orders processed by warehouses are picked manually [25]. The order picking process, a process in which humans are routed by picking lists to items’ storage locations to retrieve items for customers, is the most laborious and the most costly activity (up to 55 % of cost of the process [26]) in a typical manual assembly process. Since walking presents up to 50 % of the total time and lifting is most ergonomically stressful the logical way of improving lifting procedure (walking, turning, lifting, carrying, etc.) is to study time spent on procedure and try to reduce it by detailed research like we proposed. The problem definition of lifting procedure, which contains walking, turning, banding, applying force, lifting, carrying, lowering and posing is to determine if simulation tool is appropriate tool to assess time of the procedure. The object of study is lifting procedure of a box with a mass of $m = 13.5$ kg and dimensions of $A \times B \times C = 400 \times 400 \times 400$ mm (*height \times width \times depth*) from the floor to the table. In the initial state, the worker stands in front of the table, facing the table, and the box stands on the worker's left. The main task of the worker is to move towards the box, pick it up and put it on the table. We have separately performed the calculations for the time analyses with the MTM method, the new proposed time prediction model and the simulation experiments with the software tool JACK. To bring our case study closer to the industrial environment where these activities are common (lifting, carrying, etc.), we set up the laboratory experiment and performed the same sequence of movements. Our case study includes the next sequence of tasks (Fig. 2). First, the worker turns his body by 90° from its initial position. Then the worker starts walking parallel to the edge of the table towards to the box. He walks a distance of 1200 mm (Fig. 2a). The worker takes two lateral steps to achieve a gap between his feet to lift the box more easily. When the worker is in a balanced position, he bends to the box and prepares for the next move, which is reaching for the box (Fig. 2b). The worker picks up the box in the middle of the bottom edge. He grasps it with both hands and applies a force corresponding to the load. The worker lifts the box and regrasps it up again for easier carrying. After that, he straightens his body to the neutral position (Fig. 2c). Then he makes a 180° turn and walks back in a straight line. The distance he walks is 1200 mm (Fig. 2d).

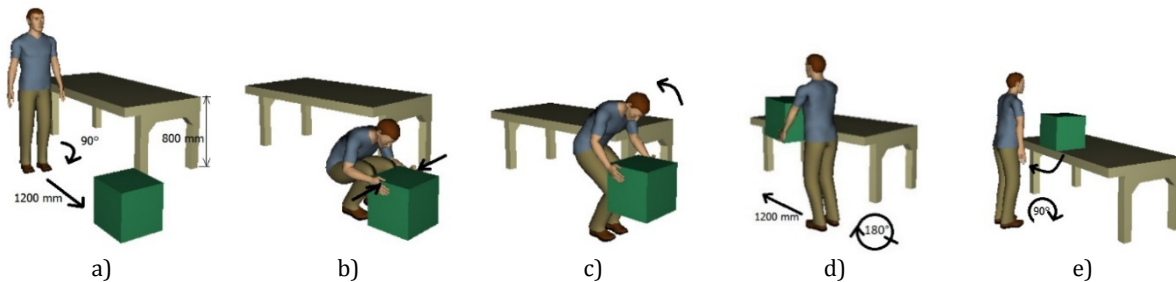


Fig. 2 Tasks from the sequence: a) turn the body by 90° and walk; b) bend the body and reach the box; c) lift the box; d) turn the body by 180° and walk; e) put the box on the table and take a neutral pose for the body

The worker turns his body towards the table. Then he puts the box on the table at a height of $h = 800$ mm. The position of the box on the table is not important and for this reason we use the "Put" movement command and not the "Position" movement command. The "Put" movement only places the box on the table without defining a tipping point, while the "Position" movement requires a specific point to be defined where the box is to be placed. After putting the box on the table, the worker returns to the neutral pose facing the table (Fig. 2e). The study was also mentioned in an already published paper [27].

MTM method

We calculated the time sequence of basic movements with the MTM method by following the instructions of Karger [28]. For our study we mainly used body, leg, and foot movements, such as sidestep, turn body, bend, arise, and walk. Other basic movements that we used emphasize hand motions, like reach, (re)grasp, apply force, move, and release. Some of these elements are used simultaneously, so we merged them together. For the time analyses of the simultaneous motions only the time for the individual motion is set so that it takes the greatest time [28].

New time prediction method (NTPM)

In our case study we used the NTPM method developed by [8] and calculated new times according to their results. The authors of the proposed NTPM took the MTM model as a basis and updated it for all movements in which the load was included. In our case we take into account the individual times from the MTM methods for the movements: turning, walking, sidestepping, bending, reaching, grasping, application of force and release. For movements where the worker handles the material (box), the exact weight of the box is taken into account, and we have calculated new individual times for the lifting, carrying, turning (carrying) and putting (lowering) movement by Eq. from 9 to 13. In addition to the weight of the box, the improved model also takes into account the angular velocity (ω_r) and the trunk-extension angle (θ_r), which we have taken from the simulation tool.

$$\omega_r(m) = -0.137662 \cdot m + 14.3881 \quad (9)$$

Empirical obtained Eq. 9 presented the correlation between the object mass m [kg], and the average angular velocity ω_r [s^{-1}]. Thereafter, the box lifting time, t_{lift} , was calculated (Eq. 10) using trunk extension angle $\Delta\theta_T$ [°] and the average angular velocity ω_r [s^{-1}].

$$t_{lift} = \frac{\Delta\theta_T}{\omega_r(m)} \quad (10)$$

Empirical obtained Eq. 11 presented the correlation between the object mass m [kg], and velocity of walking with object (carrying) v [m/s].

$$v = 5.229512605 - 0.09390347244 \cdot m \quad (11)$$

The box carrying time t_{carry} is given by Eq. 12;

$$t_{carry} = \frac{d}{v} \quad (12)$$

where d [mm] is the carrying distance, and v [m/s] is the carrying velocity in Eq. 11.

Average time for box lowering is 13 % less than that lifting under the same conditions. Thus, using the change in trunk angle $\Delta\theta_T$ [°] and angular velocity for lowering ω_r [s⁻¹], the lowering time t_{lower} was calculated as Eq. 13:

$$t_{lower} = \frac{\Delta\theta_T}{1.13 \cdot \omega_r(m)} \quad (13)$$

Simulation tool Jack

We programmed the same lifting procedure in the software tool JACK with the use of the TSB (Task Simulation Builder) tool in a virtual environment. The TSB tool automatically generates the timeline for each movement separately. We used eight different movements (go, pose, get, apply force, reach, regrasp, position, put), some of which (e.g., go) were repeated several times. Even though Jack (TSB) uses the MTM method as a basis, it makes sense to make a time comparison between TBS and MTM results, because in a virtual environment we place an avatar in different postures, which leads to a change in duration for each task.

Although the names of the individual movements differ in the MTM method and in the simulation experiment, their meaning is the same, so we can compare them. For example, it can be stated that the movement "go" in the simulation tool is used to walk, carry and turn the body, and the movement "get" is represented by the MTM method as bending and reach.

Practical experiment

In the laboratory environment we set up a system for conducting a practical experiment (Fig. 3). The experiment includes 10 healthy students. Their mean (SD) anthropometric data were: age 26 (1.5) years; body mass 75.5 (6.3) kg; height 1760 (55) mm.

All the subjects were recruited on a voluntary basis. All subjects filled in a screening questionnaire to ensure that they were in good health (i.e. that they did not suffer from any of the following: chronic illness, heart condition, musculoskeletal disorders, or injuries). They all signed a consent form approved by the institutional review board of the Faculty of Mechanical Engineering, University of Ljubljana. The experiment was conducted 20 times, two times per student. They performed the following movements: turning and walking from a starting point to the box that is 1200 mm away from the table (Fig. 3a), sidestepping, bending to the box and reaching for it (Fig. 3b), applying a force, lifting the box (Fig. 3c), turning and walking to the table (height 800 mm) (Fig. 3d), putting the box on the table and posing in the neutral position (Fig. 3e). All 20 experiments were recorded for further analyses.

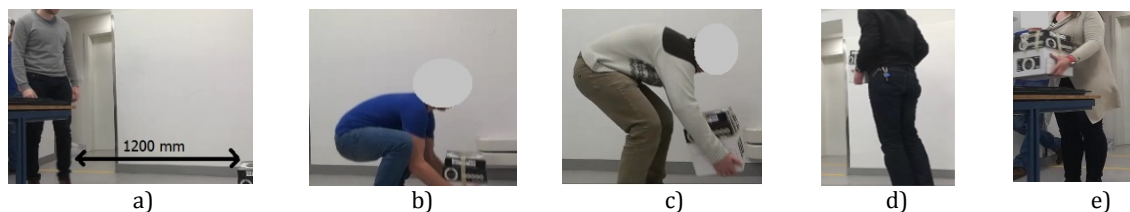


Fig. 3 Practical experiment: a) turning the body by 90° and walking; b) bending the body and reaching for the box; c) lifting the box; d) turning the body by 180° and walking; e) putting the box on the table and taking a neutral pose of the body

2.3 Comparison of the method

All methods used in our research are based on the MTM methodology, with the exception of the practical experiment, which is an indicator of the actual performance of the subjects. As shown in Table 1, different terms are used for the basic movements in all four methods. We can see that the main difference is only in the naming and the merging of several movements together; the purposes and meanings are the same everywhere, so in summary we can say that the methods are comparable. The results of each method were obtained in different ways: for the MTM classical method, the results were calculated using tables and recommendations [28]. In the Jack simulation environment, a time report is automatically generated based on the avatar, its trajectories, and the parameters you insert in the program. In the NTPM method, the results were obtained by calculation according to the classical MTM method and the addition of taking into account the mass according to Eq. from 9 to 13.

Table 1 Comparison and description of the basic movements for all four methods

MTM method	NTPM	Jack simulation	Practical experiment
Turn body by 90°	Same as MTM	Go (turn)	Turn
Walk	Same as MTM	Go	Walk
Left sidestep	Same as MTM	Pose	Sidestep
Right sidestep	Same as MTM		
Bend	Same as MTM	Get	Bend and reach
Reach	Same as MTM		
Grasp	Same as MTM		
Apply force	Same as MTM	Wait (time from MTM)	Apply force
Arise	Lift	Position (arise and reach)	
Move the box		Lift	
Regrasp		Regrasp	
Turn body by 90°	Carry	Go (turn)	Turn by 180°
Turn body by 90°		Go	Walk
Walk	Carry	Go (turn)	Turn
Turn body by 90°	Carry	Put	Put
Moving the box	Lower		
Release the box	Same as MTM	Pose	Neutral pose
Neutral pose	Same as MTM		

2.4 Parameter combination

We performed 25 simulation experiments (5 different masses of the box multiplied by 5 different dimensions of the box) with different combinations of masses and dimensions of the box (parameters in Fig. 1). We compared all versions using the MTM method, the NTPM and JACK the simulation tool, focusing on the total time.

3. Results and discussion

In this section we present the results of the time analyses for the job and task. The section is divided into two parts. The first part shows the results of the time analyses of the individual tasks, which were achieved using all four methods. The second part presents the results of the influential parameters obtained with all three methods: MTM, NTPM, and Jack simulation tool.

3.1 Time analyses of the individual tasks

All movements in all methods were unified so that they are comparable with each other. In the MTM method we combined simultaneous movements so that we obtained a total of 11 movements which are the same in all four methods we compared. The movements and the corresponding times are shown in Table 2. The box we used in this part of the study has a dimension of 400 mm and weighs 13.5 kg.

Table 2 Comparison of the individual times of the basis movement for all four methods

#	Task description	MTM Time [s]	NTPD Time [s]	Jack Time [s]	Experiment Time [s]
1	Turn body by 90°	0.67	0.67	0.2	0.97
2	Walk	0.76	0.76	1.34	1.97
3	Sidestep	0.64	0.64	0.64	0.78
4	Bend, reach and grasp the box	1.30	1.30	2.54	1.02
5	Apply force	0.38	0.38	0.38	0.59
6	Lift the box	1.38	3.50	3.16	1.55
7	Turn by 180°	2.01	0.39	0.78	1.47
8	Walk to the table	0.67	1.10	0.95	1.37
9	Turn by 90°	1.34	0.18	0.17	0.61
10	Put the box on the table	0.35	3.10	1.76	1.25
11	Neutral pose of the body	0.52	0.52	0.50	0.87
	Total time TT [s]	10.12	12.54	12.42	12.45

The total time TT [s] for each method was calculated by Eq. 1. It can be noticed that the terms of the Eq. 1 and task names are not the same. These are only word differences and not substantive ones. Here are the explanations of meaning for each terms of Eq. 1 according to task description. Task #1 “turn body by 90°”, #2 “walk” and #3 “sidestep” contributes time to the first term of Eq. 1 “ t_{walk} ” and has a functional dependence determined by Eq. 2. Task “band, reach and grasp” contributes time t_{band} to the total time and is defined by Eq. 3. Task #5 “apply force” is described by term t_{apply} and Eq. 4. Task #5 “lift the box” contributes to total time the time t_{lift} and its functionality is described by Eq. 5. Task numbered from 7 to 9 means carrying (walking with object) and its contributions are defined by t_{carry} and functional dependencies by Eq. 6. Task #10 “put the box on the table” means to lower the load (object) and is defined by term t_{lower} and Eq. 7. The last task “neutral pose of the body” contributes time to term t_{pose} and is described by Eq. 8.

The values for the times of all movements are not completely consistent when comparing all four methods. It is clear that the different movements contribute different proportions to the total time. Therefore, the same movement has a greater or less strong effect on the total time if we use different methods. We can see that the movements turning the body and walking, which were obtained by the experimental study, contribute a relatively large proportion to the total time compared to the other methods. The experimental study was conducted in a laboratory environment. The subjects were inexperienced and were only instructed to perform the operations at a pace that would not make them feel uncomfortable if the sequence of tasks was repeated over a long period of time. If we exclude these factors, the walking time would be closer to the results of the other methods.

The next comparison refers to the results obtained with the simulation tool Jack and the NTPM. In the NTPM we only focus on the following movements: lift, turn by 180° with object (carry), walk with object (carry), turn by 90° (carry), and put (lower) – task numbers from 6 to 10, because only these movements take into account the weight of the load and are upgraded by the MTM method. We calculated times by equations from 9 to 13 with the following parameters: $m = 13.5$ kg; $d_{walking} = 1200$ mm; $d_{turning} = 200$ mm; and $\Delta\theta_T = 43.6^\circ$. It can be seen that the values determined with these two methods for the times of the movements “lift”, and “put” (lower) are longer than the times determined with the other two methods (MTM and experimental study). The reason for this difference is that both methods (NTPM, Jack tool) take into account the trunk-extension angle, which also increases the value of time and ensures that the worker has enough time to lift and lower the loads in an ergonomically correct way.

The results of the MTM method indicate that the worker spends most of the time on body turns compared to the other methods. Both in the simulation experiment with the Jack tool and in the experimental study (assuming that the test subjects would be trained) less time is needed to turn the body, because in reality we combine rotation and walking into a simultaneous movement for which we need less time.

Jack tool compared to other methods spend least time for task “turn the body”. The reason is that Jack doesn't have a separate “Turn” movement, but you have to use a “Go” movement. This movement requires moving the avatar by a certain distance, and not just turning the whole body in the same place. So we used a distance of 200 mm for the turn and a predetermined speed (specified in the simulation tool) and thus obtained the task time. For turning the body you spend more time than for walking, because of the higher complexity of the kinematics of the joints and here is the time difference. The second task, which is at contrary longer in comparison with other method, is “band, reach and grasp” task. There are two reasons for this. First is that it is necessary to place the avatar in a certain distance from object and also to “attach” the hand for “grasp” movement to a certain place, which, however, differs in each simulation and affect the task's time. The repeatability of this task is achieved only through experience and long-term use of simulation tool. The second reason is that the task itself is predefined in the simulation environment so that the avatar leans towards the object with stretched legs, which did not suit us in terms of comparison with other methods. So we subsequently changed the angles of certain joints (knees, hips, and torso) and thus gained time for this task. When changing such a complex task, the simulation tool determine longer time.

The total time for all the tasks obtained with the MTM method is $t = 10.12$ s; using the Jack simulation tool it is $t = 12.42$ s; for the practical experiment it is $t = 12.45$ s; and the total time obtained with the NTPM method is $t = 12.54$ s.

In summary, we can say that all methods, except MTM, give almost the same result for the total job time and that we can use any method to design a new job with a sequence of basic movements. In many cases PMTS (the MTM method) predicts a shorter time than the worker is able to work without injury after some time. In such a case, this statement proves to be correct in comparison with the other methods, as the MTM method predicts 20 % less time to perform the same sequence of movements.

In industry it is necessary to plan and design work processes and sequences of jobs as quickly as possible, but it is also necessary to adapt the tasks to the workers and to think about their well-being. With the Jack simulation tool, it is easy and quick to check the ergonomics and suitability of the jobs for workers using "what if" scenarios. This option allows us to modify the various parameters of the simulated objects (weight, dimension, shape) and avatars (gender, anthropometric characteristics). From this part of our study we can conclude that the Jack simulation tool is a suitable method for designing jobs and workplaces when only the total time and not the individual times are of interest. If we want to optimize the work cycle for the study we are discussing, it is necessary to look deeper into the individual tasks and find the relationships between the values of the individual times and the individual tasks, and determine which characteristics (e.g. steps, turning procedure, bending procedure, trunk angles, etc.) have the greatest influence on a particular method.

3.2 Influential parameters

Tables 3 to 5 show the results by MTM, NTPM and Jack simulation. The aim of the influential parameter study was to determine how the individual method takes into account the weight parameter and the load size parameter during the lifting procedure. The results show the total times of the entire sequence of all tasks. The load that appeared in the study is a cubic box without handles. The mass of the box varies from 4.5 kg to 20.5 kg. The dimensions (height = width = depth) of the box vary from 200 mm to 600 mm. Therefore, the study presents 25 combinations of mass and dimensional parameters.

MTM method

When using the MTM method, the mass and dimensions of the box are taken into account when calculating the total time (Table 3). The total time of the tasks and the mass of the box are linearly dependent ($R^2 = 0.9998$; Fig. 4a). The same applies to the dimensions ($R^2 = 1$; Fig. 4b). Therefore, if we use the MTM method, we can predict with reasonable certainty the total time for this sequence of tasks for each size and mass of the box. The problem with prediction is that these times are often too short, which means that a worker is not able to perform all movements without further discomfort.

Table 3 MTM: total time analysis for different dimensions and masses of the box

		Mass of the box [kg]					Time [s]
		4.5	9.0	13.5	18.0	20.5	
Dimension of the box [mm]	200	9.32	9.60	9.89	10.17	10.34	
	300	9.43	9.72	10.01	10.29	10.44	
	400	9.54	9.83	10.12	10.42	10.56	
	500	9.66	9.95	10.24	10.54	10.68	
	600	9.77	10.06	10.36	10.66	10.80	

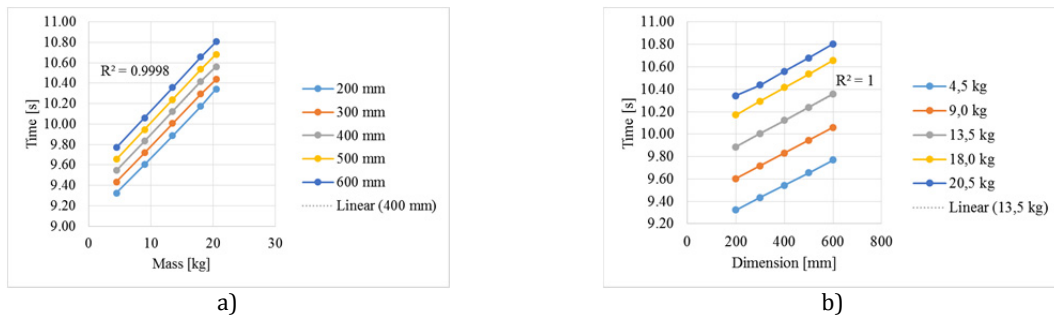


Fig. 4 Relationship between the total time and the parameters: a) mass, and b) dimension using the MTM method

NTPM method

Table 4 shows the values of the total times obtained using the NTPM method. It can be seen immediately that the dimensions of the box are not included (Fig. 5b). The relationship between the weight and the total time is linear ($R^2 = 0.9977$, Fig. 5a), so for this sequence of movements we can predict the value of the total time for each mass of the box. To calculate the results according to this method we use several parameters (angular velocity, trunk-extension angle) and the mass of the box are used, which gives more realistic total times compared to the MTM method.

Table 4 NTPM: total time analyses for different dimensions and masses of the box

		Mass of the box [kg]					Time [s]
		4.5	9.0	13.5	18.0	20.5	
Dimension of the box [mm]	Not included		11.34	11.88	12.54	13.13	13.53
			11.34	11.88	12.54	13.13	13.53
			11.34	11.88	12.54	13.13	13.53
			11.34	11.88	12.54	13.13	13.53
			11.34	11.88	12.54	13.13	13.53

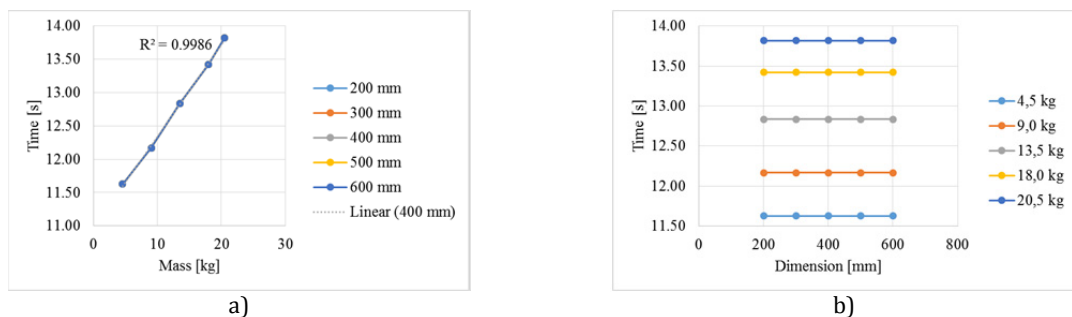


Fig. 5 Relationship between the total time and the parameters using the NTPM: a) mass, and b) dimension

Jack simulation tool

The results in Table 5 show that the time changes minimally when the mass of the load changes. Therefore, the mass of the box has minimal effect on the value of total time, which is not logical or empirical, as shown in Fig. 6a – relationship between total time and the mass parameter.

Table 5 Jack simulation tool: total time analysis for different dimensions and masses of the box

		Mass of the box [kg]					Time [s]
		4.5	9.0	13.5	18.0	20.5	
Dimension of the box [mm]	200	12.06	12.20	12.34	12.51	12.57	
	300	12.18	12.31	12.43	12.56	12.62	
	400	12.15	12.28	12.42	12.54	12.60	
	500	12.15	12.27	12.40	12.53	12.57	
	600	12.36	12.48	12.60	12.72	12.79	

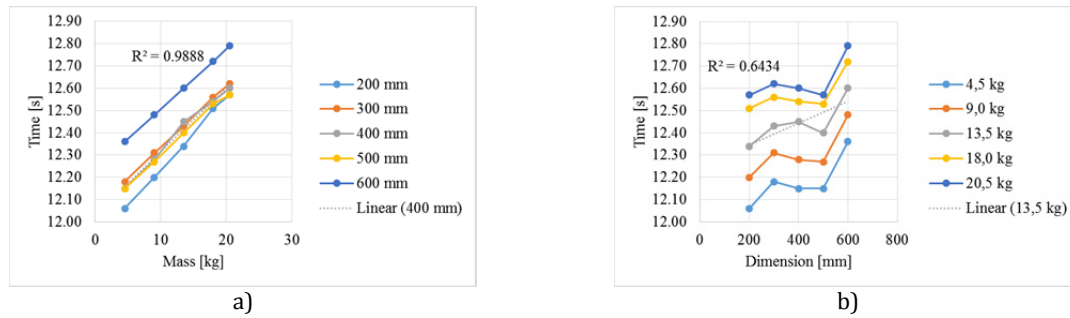


Fig. 6 Relationship between the total time and the parameters using the Jack simulation tool: a) mass, and b) dimension

From the results of the analysis of the dimensions of the box it is clear that the dimensions affect the lengthening/shortening of the total time. The change of the duration of a task happens because it is necessary to adjust all positions of the body and joints separately for each task in order to handle the different sizes of the boxes. The total time varies because we cannot adjust the joints in exactly the same way, which is similar to the situation in reality (Fig. 6b).

We can conclude that the results of the simulation show that the mass of the box has a limited effect on the total time. However, a more detailed analysis shows that the change in mass strongly affects the ergonomic analysis of the program package (Lower-Back Analysis – Fig. 7, Manual Handling Limits, Metabolic Energy Expenditure, NIOSH Lifting Analysis etc.). The result on Fig. 7 shows force (L4/L5 force) affecting the lumbar vertebrae 4 and 5. If we want to obtain a "correct" correlation between the mass and the value of the total time in a simulation tool, we should try to place the avatar in different positions and, based on the ergonomic results, extend/shorten the time of each task and consequently obtain more regular values for the total times.

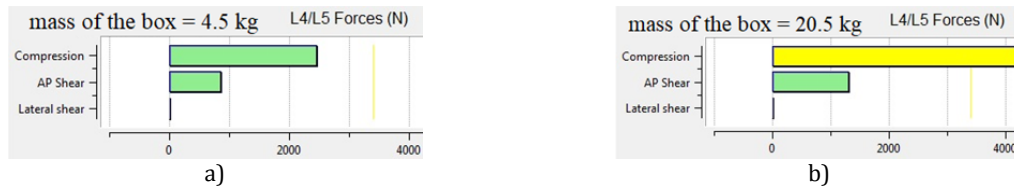


Fig. 7 Lower-back analysis – Results for L4/L5 forces [N] based on different masses of the box: a) 4.5 kg, and b) 20.5 kg

4. Conclusion

The simulation model of the lifting procedure is developed for prediction of the total time required for basic manual assembly tasks with different load parameters. The results of the simulation model were compared with different time prediction methods (MTM, NTPM) and verified by laboratory experiment. Workers in manual assembly suffer from rigorous time standards. With the implementation of a simulation model, the time standards of the task can be efficiently redefined to make it worker-friendly.

Our research can be divided into two parts. The first part concerned the individual times and the total times determined with four methods, two of which are time prediction methods (MTM, NTPM), another one is simulation of the worker obtained by Siemens Jack simulation tool, and the last one is a laboratory experiment. The research was carried out on basic manual tasks, which cover only a part of the MMH tasks. The main objective was to determine whether the simulation tool is suitable not only for ergonomic analyses but also for setting time standards for the workers. The literature shows that workers under traditional standards do not have enough time to perform basic tasks, resulting in work-related musculoskeletal disorders and injuries. As we found out in the case study, the basis is the classical MTM method, which gives the shortest total time of the lifting sequence. The NTPM upgrades the MTM method by extending the times due to the load consideration, which results in the worker having more time to recover. The simulation tool, which is also based on the MTM method, leads to differences in the total time value in comparison to the MTM method. The reason is that the TBS tool also takes into account the trajectories of the body, arms, legs, and the way the avatar grasp the load and adds addition-

al time on these tasks compared to MTM times. We also verified the sequence of the tasks with an experiment that showed very similar result as the simulation tool. From this we can conclude that simulation is a suitable tool for designing time standards for a sequence of basic methods, although the duration of a single task varies depending on the method used.

The second part of the study focuses on the load of the lifting procedure. We studied the influence of the load on the total time of the lifting procedure. Parameters of the load that were studied in more detail are the mass of the load and the dimensions of the load. As load we used a cube-shaped box with an evenly distributed weight without handles. The case study was performed as a sequence of tasks with 25 different load parameter combinations (5 different box masses x 5 different box dimensions). The case study was performed using three different methods, MTM, NTPM, and the Siemens Jack simulation tool. Depending on the method used, we obtained different results. We found that the MTM method takes both parameters into account but, as reported above, gives too short times for a healthy working practice. The NTPM method only considers the mass of the box, while the dimensions are not considered. The simulation tool, on the other hand, takes the dimensions of the box into account, but the mass of the box has minimal effect on the duration of the tasks, which is in contrast to practice. To overcome this drawback, we have found that the mass has a "correct" effect on the ergonomic result shown by the Lower Back Analysis, so the more detailed study of ergonomics implemented in the simulation tool can lead to a better correlation between the time of execution of each task and the mass of the load.

Further research work will include a determination of this correlation, and how this can be used for an even more realistic estimation of the total time.

Acknowledgement

The work was carried out in the framework of the GOSTOP programme (OP20.00361), which is partially financed by the Republic of Slovenia – Ministry of Education, Science and Sport, and the European Union – European Regional Development Fund. The authors would also like to thank the Research and Development Agency (ARRS) of the Republic of Slovenia, which made this study possible.

References

- [1] Nogueira, H.C., Locks, F., Barbieri, D.F., Oliveira, A.B. (2018). How does the biomechanical exposure of the upper body in manual box handling differ from exposure in other tasks in the real industrial context?, *International Journal of Industrial Ergonomics*, Vol. 68, 8-14, doi: [10.1016/j.ergon.2018.05.015](https://doi.org/10.1016/j.ergon.2018.05.015).
- [2] Borgs, S.P., La Delfa, N.J., Dickerson, C.R. (2019). An evaluation of off-axis manual forces and upper extremity joint moments during unilateral pushing and pulling exertions, *Ergonomics*, Vol. 62, No. 1, 52-64, doi: [10.1080/00140139.2018.1525501](https://doi.org/10.1080/00140139.2018.1525501).
- [3] Leber, M., Bastič, M., Moody, L., Schmidt Krajnc, M. (2018). A study of the impact of ergonomically designed workplaces on employee productivity, *Advances in Production Engineering & Management*, Vol 13, No. 1, 107-117, doi: [10.14743/apem2018.1.277](https://doi.org/10.14743/apem2018.1.277).
- [4] Rasmussen, C.D.N., Højberg, H., Bengtsen, E., Jørgensen, M.B. (2018). Identifying knowledge gaps between practice and research for implementation components of sustainable interventions to improve the working environment – A rapid review, *Applied Ergonomics*, Vol. 67, 178-192, doi: [10.1016/j.apergo.2017.09.014](https://doi.org/10.1016/j.apergo.2017.09.014).
- [5] Dianat, I., Molenbroek, J., Castellucci, H.I. (2018). A review of the methodology and applications of anthropometry in ergonomics and product design, *Ergonomics*, Vol. 61, No. 12, 1696-1720, doi: [10.1080/00140139.2018.1502817](https://doi.org/10.1080/00140139.2018.1502817).
- [6] de Mattos, D.L., Ariento Neto, R., Merino, E.A.D, Forcellini, F.A. (2019). Simulating the influence of physical overload on assembly line performance: A case study in an automotive electrical component plant, *Applied Ergonomics*, Vol. 79, 107-121, doi: [10.1016/j.apergo.2018.08.001](https://doi.org/10.1016/j.apergo.2018.08.001).
- [7] Lanzotti, A., Vanacore, A., Tarallo, A., Nathan-Roberts, D., Coccorese, D., Minopoli, V., Carbone, F., d'Angelo, R., Grasso, C., Di Gironimo, G., Papa, S. (2019). Interactive tools for safety 4.0: Virtual ergonomics and serious games in real working contexts, *Ergonomics*, Vol. 63, No. 3, 324-333, doi: [10.1080/00140139.2019.1683603](https://doi.org/10.1080/00140139.2019.1683603).
- [8] Harari, Y., Bechar, A., Raschke, U., Riemer, R. (2017). Automated simulation-based workplace design that considers ergonomics and productivity, *International Journal of Simulation Modelling*, Vol. 16, No. 1, 5-18, doi: [10.2507/IJSIMM16\(1\)1.355](https://doi.org/10.2507/IJSIMM16(1)1.355).
- [9] Harari, Y., Riemer, R., Bechar, A. (2018). Factors determining workers' pace while conducting continuous sequential lifting, carrying, and lowering tasks, *Applied Ergonomics*, Vol. 67, 61-70, doi: [10.1016/j.apergo.2017.09.003](https://doi.org/10.1016/j.apergo.2017.09.003).

- [10] Gupta, I., Kalra, P., Chawla, P., Singh, J. (2018). Evaluation of pilot's seat design of civil aircraft for Indian anthropometric data by using Delmia human software, *Procedia Manufacturing*, Vol. 26, 70-75, doi: [10.1016/j.promfg.2018.07.009](https://doi.org/10.1016/j.promfg.2018.07.009).
- [11] Padula, R.S., Coury, H.J.C.G. (2003). Sagittal trunk movements during load carrying activities: A pilot study, *International Journal of Industrial Ergonomics*, Vol. 32, No. 3, 181-188, doi: [10.1016/S0169-8141\(03\)00062-3](https://doi.org/10.1016/S0169-8141(03)00062-3).
- [12] Martinez, R., Bouffard, J., Michaud, B., Plamondon, A., Côté, J.N., Begon, M. (2019). Sex differences in upper limb 3D joint contributions during a lifting task, *Ergonomics*, Vol. 62, No. 5, 682-693, doi: [10.1080/00140139.2019.1571245](https://doi.org/10.1080/00140139.2019.1571245).
- [13] Corbeil, P., Plamondon, A., Handrigan, G., Vallée-Marcotte, J., Laurendeau, S., Ten Have, J., Manzerolle, N. (2019). Biomechanical analysis of manual material handling movement in healthy weight and obese workers, *Applied Ergonomics*, Vol. 74, 124-133, doi: [10.1016/j.apergo.2018.08.018](https://doi.org/10.1016/j.apergo.2018.08.018).
- [14] Parida, R., Ray, P.K. (2015). Biomechanical modelling of manual material handling tasks: A comprehensive review, *Procedia Manufacturing*, Vol. 3, 4598-4605, doi: [10.1016/j.promfg.2015.07.539](https://doi.org/10.1016/j.promfg.2015.07.539).
- [15] Plamondon, A., Larivière, C., Denis, D., Mecheri, H., Nastasia, I. (2017). Difference between male and female workers lifting the same relative load when palletizing boxes, *Applied Ergonomics*, Vol. 60, 93-102, doi: [10.1016/j.apergo.2016.10.014](https://doi.org/10.1016/j.apergo.2016.10.014).
- [16] Theurel, J., Desbrosses, K., Roux, T., Savescu, A. (2018). Physiological consequences of using an upper limb exoskeleton during manual handling tasks, *Applied Ergonomics*, Vol. 67, 211-217, doi: [10.1016/j.apergo.2017.10.008](https://doi.org/10.1016/j.apergo.2017.10.008).
- [17] Pinder, A.D.J., Boocock, M.G. (2014). Prediction of the maximum acceptable weight of lift from the frequency of lift, *International Journal of Industrial Ergonomics*, Vol. 44, No. 2, 225-237, doi: [10.1016/j.ergon.2012.11.005](https://doi.org/10.1016/j.ergon.2012.11.005).
- [18] Lee, T.-H. (2003). Minimal acceptable handling time intervals for lifting and lowering tasks, *Applied Ergonomics*, Vol. 34, No. 6, 629-634, doi: [10.1016/S0003-6870\(03\)00050-4](https://doi.org/10.1016/S0003-6870(03)00050-4).
- [19] Straker, L.M., Stevenson, M.G., Twomey, L.T. (1996). A comparison of risk assessment of single and combination manual handling tasks: 1. Maximum acceptable weight measures, *Ergonomics*, Vol. 39, No. 1, 128-140, doi: [10.1080/00140139608964439](https://doi.org/10.1080/00140139608964439).
- [19] Straker, L.M., Stevenson, M.G., Twomey, L.T., Smith, L.M. (1997). A comparison of the risk assessment of single and combined manual handling tasks: 3. Biomechanical measures, *Ergonomics*, Vol. 40, No. 7, 708-728, doi: [10.1080/001401397187856](https://doi.org/10.1080/001401397187856).
- [21] Ahmed, S., Gawand, M.S., Irshad, L., Demirel, H.O. (2018). Exploring the design space using a surrogate model approach with digital human modelling simulation, In: *Proceedings of the ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Volume 1B: 38th Computers and Information in Engineering Conference*, Quebec, Canada, 1-13, doi: [10.1115/DETC2018-86323](https://doi.org/10.1115/DETC2018-86323).
- [22] Vujica Herzog, N., Harih, G. (2019). Decision support system for designing and assigning ergonomic workplaces to workers with disabilities, *Ergonomics*, Vol. 63, No. 2, 225-236, doi: [10.1080/00140139.2019.1686658](https://doi.org/10.1080/00140139.2019.1686658).
- [23] Harari, Y., Bechar, A., Riemer, R. (2020). Workers' biomechanical loads and kinematics during multiple-task manual material handling, *Applied Ergonomics*, Vol. 83, Article No. 102985, doi: [10.1016/j.apergo.2019.102985](https://doi.org/10.1016/j.apergo.2019.102985).
- [24] Harari, Y., Bechar, A., Riemer, R. (2019). Simulation-based optimization methodology for a manual material handling task design that maximizes productivity while considering ergonomic constraints, *IEEE Transactions on Human-Machine Systems*, Vol. 49, No. 5, 440-448, doi: [10.1109/THMS.2019.2900294](https://doi.org/10.1109/THMS.2019.2900294).
- [25] Vujica Herzog, N., Buchmeister, B., Beharic, A., Gajsek, B. (2018). Visual and optometric issues with smart glasses in Industry 4.0 working environment, *Advances in Production Engineering & Management*, Vol. 13, No. 4, 417-428, doi: [10.14743/apem2018.4.300](https://doi.org/10.14743/apem2018.4.300).
- [26] Klodawski, M., Jachimowski, R., Jacyna-Golda, I., Izdebski, M. (2018). Simulation analysis of order picking efficiency with congestion situations, *International Journal of Simulation Modelling*, Vol. 17, No. 3, 431-443, doi: [10.2507/IJSIMM17\(3\)438](https://doi.org/10.2507/IJSIMM17(3)438).
- [27] Turk, M., Resman, M., Herakovič, N. (2018). Preparation of papers for IFAC conferences & symposia: Computer-aided processing of manual assembly operations with integration of simulation tools in production processes, *IFAC-PapersOnLine*, Vol. 51, No. 2, 813-818, doi: [10.1016/j.ifacol.2018.04.014](https://doi.org/10.1016/j.ifacol.2018.04.014).
- [28] Karger, D.W., Bayha, F.H. (1961). *Engineered Work Measurement*, First edition, Industrial Press, New York, USA.

Manufacturer's customer satisfaction incentive plan for duopoly retailers with Cournot or collusion games

Hu, H.^{a,b,*}, Zhang, Z.^a, Wu, Q.^a, Han, S.^a

^aSchool of Economics and Management, Yanshan University, Qinhuangdao, P.R. China

^bResearch Center for Regional Economic Development, Yanshan University, Qinhuangdao, P.R. China

ABSTRACT

To increase customer satisfaction (CS) which is closely linked to corporate reputation, revenue and customer loyalty, manufacturers will provide incentives to retailers in supply chain management. This paper focuses on two types of incentives that a manufacturer may provide to retailers: customer satisfaction index bonus (CSI bonus) and customer satisfaction assistance, and studies the optimal customer satisfaction incentive plan of the manufacturer when duopoly retailers adopt Cournot or collusion game. By comparing the equilibrium of the two games, we conducted a preference analysis of both the manufacturer and the retailers. The results showed that no matter what kind of games the duopoly retailers take, the manufacturer will provide customer satisfaction assistance to the retailers to increase the customer satisfaction. However, if the duopoly retailers take Cournot behaviour, only when the wholesale price is greater than the threshold, the manufacturer will provide the retailers with customer satisfaction index bonus. The manufacturer always prefers to Cournot behaviour, and the retailers always prefer to collusion behaviour. In addition, this paper also investigated the effect of customer satisfaction incentives on the manufacturer, and found that it will help the manufacturer obtain more demand and higher profits.

© 2020 CPE, University of Maribor. All rights reserved.

ARTICLE INFO

Keywords:

Supply chain management;
Manufacturer's incentive plan;
Customer satisfaction;
Duopoly retailers;
Game;
Cournot game;
Collusion

*Corresponding author:

huhaiju@ysu.edu.cn
(Hu, H.)

Article history:

Received 12 August 2020
Revised 23 September 2020
Accepted 29 September 2020

1. Introduction

Nowadays, CS is concerned by various industries, so it is imperative to improve it. CS plays such an importance role because it is closely related to a company's reputation, marketing activities and revenues [1-2]. High CS will result in more repeat purchases, cross purchases and a good reputation [3-5]. When CS is low, it will make consumers have a bad impression of the company, which will affect the reputation of the company and make it difficult for the company to last for a long-term development. Therefore, to encourage retailers to improve CS, many manufacturers implement some form of supply chain incentive measures. In practice, manufacturers typically offer two types of CS incentives to retailers: CSI bonus and CS assistance [6-7]. Specifically, CSI bonuses are one-time incentives given by manufacturers to retailers based on CSI scores. According to the 2017 CSI automotive industry research results in China, the CSI scores of Audi (86.6), BMW (80.1) and Mercedes-Benz (71.5) are the top three in the luxury car series; Toyota (81.8), Chevrolet (77.8) and Volkswagen (73.1) are listed as the top three in the mainstream car series. They gave retailers a certain CSI reward based on CSI scores. CS assistance is an investment made by manufacturers to share the cost of retailers' efforts to improve CS. For example, Ford and Lexus train the employees of investment retailers for free to reduce the input cost of retailers' CS efforts; Benz and Saturn offer free consulting services to retailers to improve CS.

CS was first proposed by Cardozo [8], who argued that improving CS would lead to a repurchasing behaviour. Wang and Zhao [1] established a multilayer linear model discussed the impact of CS on the value of shareholder. There are also many studies to improve CS by implementing some kinds of incentive. The CSI was first established by Sweden in 1989. It is an index calculated based on customer evaluation of enterprise product and service quality. Hauser *et al.* [9] were the first to study CSI bonuses, and they considered a two-phase supply chain model, in which the CS incentive plan increased the future demand of the second phase. Based on this, Chu and Desai [6] proposed another incentive: CS assistance, which encourages retailers to improve CS, and they found that the manufacturer would provide these two incentives in their CS incentive plan when the actors all have strong pricing power. Wang *et al.* [7] studied the optimal CS incentive plan for manufacturers in the four supply chain models, and they further found that when the manufacturer's market pricing ability is weak and the wholesale price is lower than a certain value, the manufacturer only provides CS assistance to the retailers.

These literatures mainly study that a manufacturer provide incentives to a retailer to improve CS, but the situation of providing incentives to duopoly retailers has not been studied. Generally speaking, there are many duopoly supply chains, such as offline sales Gome and Suning, Wal-Mart and Tesco, and online sales such as JD.com and Tmall. By cooperating with duopoly retailers, the manufacturer can quickly expand their market influence, and it must also improve CS in the face of fierce competition between similar products, so as to win in the competition. There are different forms of competition among duopoly retailers. Competition between duopoly retailers is a common phenomenon in various industries. They (e.g. Wal-Mart and Tesco) determine their own selling price and order quantity respectively and pursue their own profit maximization. Cooperation is another behaviour between duopoly retailers, who jointly set their selling price so that the entire downstream supply chain of profit maximization. Yang and Zhou [10] studied the three behaviours of the duopoly retailers Cournot, collusion and Stackelberg, and conducted preference analysis based on the best decision. They finally found that if the duopoly retailers take collusion behaviour, its selling price is the highest; However, if the duopoly retailers take Cournot behaviour, its selling price is the lowest. Huang *et al.* [11] established six decentralized models and a centralized model according to different rights structures. Modak *et al.* [12] studied the Cournot and collusion model of duopoly retailers in a closed-loop supply chain. Herbon [13] studied the price, profit and market share of two competition retailers in the context of consumer information asymmetry. Li *et al.* [14] studied a combination of strategies of one manufacturer and multiple suppliers for cooperative supply. Lang and Shao [15] studied the product portfolio decisions of two retailers under competition situations based on the MNL model. Chai *et al.* [16] studied the price competition of two competing retailers under stochastic demand disruption.

In supply chain coordination, there are many coordination contracts, such as volume discount contracts [17], two-part tariff contracts [18-19], revenue-sharing contracts [20-22], incentive adjustment policy [23], cost sharing contracts [24], etc. Hu *et al.* [25] studied the competition among DCSC members and the optimal strategy in these two coordination situations. This paper focuses on improving supply chain performance through incentive policies. Cheng *et al.* [26] explored incentive contracts under conditions of uncertain demand and asymmetric information. Deng *et al.* [27] studied the incentive mechanism of CS in the distributed service chain. Hu *et al.* [28] studied the retailers provide incentives for manufacturers to improve product quality, thereby increasing consumers' willingness to purchase. Deng *et al.* [29] designed an effective CS incentive contract taking budget constraints into account, and they found that even in a single period environment that the customers would not bring in future business, CS incentives benefit the manufacturer.

This paper builds 2 kinds of game scenarios of the duopoly retailers. The specific arrangements are as follows: Section 2 describes the model. Section 3 studies the manufacturer's CS incentive plan in two game scenarios, and compares and analyses the decision results. Section 4 studies the impact of CS incentives on the manufacturer. Section 5 summarizes conclusions, points out deficiencies and looks to the future.

2. Model description

This study focuses on the supply chain consisting of one manufacturer and two retailers. The manufacturer wholesales the same product to the two retailers at the same price, and then retailers sell the product to consumers in the same market.

The demand function is based on Chu and Desai [6] which merges short-term demand and long-term demand into a single demand function, and add to this demand function the effect of another retailer's price, resulting in a new demand function as follows.

$$D_i = \alpha_i - p_i + b_i + \beta p_{(3-i)} \quad (i = 1,2) \tag{1}$$

D_i : Retailer i 's demand function;

α_i : Retailer i 's market potential (initial reputation);

p_i : Retailer i 's selling price;

b_i : Retailer i 's CS efforts;

β : The degree of substitutability between retailers, reflects the impact of retailer's marketing mix decisions on customer demand, with $0 \leq \beta < 1$.

Assume retailer i 's CS effort cost is b_i^2 . CS incentive plan consists of two parts: (1) CSI bonus, the bonus per unit CS effort is η , CSI bonus is measured by retailer i 's CS effort b_i direct measurement, the manufacturer gives the retailer i 's the total amount of CSI bonus is ηb_i . (2) CS assistance, the assistance per unit CS effort is x_b , so the retailer i 's CS effort cost reduction $b_i x_b$, cost is $b_i(b_i - x_b)$, then the cost incurred by the manufacturer is x_b^2 .

Retailer i the profit function π_{ri} is:

$$\pi_{ri} = (p_i - w_m)D_i - b_i(b_i - x_b) + \eta b_i \tag{2}$$

Where w_m is the wholesale prices offered by the manufacturer.

Assume that the manufacturer's manufacturing cost is 0, the manufacturer's profit function π_m is:

$$\pi_m = \sum_{i=1}^2 (w_m D_i - \eta b_i - x_b^2) \tag{3}$$

3. Model analysis

In the model analysis below, we assume that the manufacturer is the leader of Stackelberg and the duopoly retailers is its follower, and the retailers can play either Cournot or collusion game.

3.1 Retailers play the Cournot game

In this section, it is assumed that duopoly retailers adopt a Cournot approach in which each retailer sets its own selling price and CS effort by assuming that its competitors' selling price and CS effort.

Our game sequence is:

(1) the manufacturer determines w_m, η, x_b ;

(2) retailer i according to the given w_m, η, x_b to determine p_i, b_i .

In order to find the equilibrium of the subgame of the two-stage game, the backward induction method is adopted.

For any given w_m, η, x_b , donating $\frac{\partial \pi_{r1}}{\partial p_1} = 0, \frac{\partial \pi_{r1}}{\partial b_1} = 0$ we have the optimal selling price p_1^* and the best CS effort b_1^* ; Retailer 2 passed solve $\frac{\partial \pi_{r2}}{\partial p_2} = 0, \frac{\partial \pi_{r2}}{\partial b_2} = 0$ to determine the optimal selling price p_2^* and the optimal CS effort b_2^* . Followed, putting $p_1^*, b_1^*, p_2^*, b_2^*$ into equation (3), the manufacturer solves $\frac{\partial \pi_m}{\partial w_m} = 0, \frac{\partial \pi_m}{\partial \eta} = 0, \frac{\partial \pi_m}{\partial x_b} = 0$ to determine the optimal wholesale price w_m^* , the optimal CS effort bonus per unit η^* , the optimal CS effort assistance per unit x_b^* .

When the profit function is a concave function on its decision variable, the optimal solution can maximize the profit. Therefore, the concaveness of the profit function is tested by calculating the Hessian matrix.

As can be seen from Eq. 1, the Hessian matrix of π_{r1} is:

$$H_{c1} = \begin{bmatrix} \frac{\partial^2 \pi_{r1}}{\partial p_1^2} & \frac{\partial^2 \pi_{r1}}{\partial p_1 \partial b_1} \\ \frac{\partial^2 \pi_{r1}}{\partial b_1 \partial p_1} & \frac{\partial^2 \pi_{r1}}{\partial b_1^2} \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$$

$\Delta_1 = -2 < 0$; $\Delta_2 = 3 > 0$. Retailer 1's profit function π_{r1} is its decision variable p_1 with b_1 strict concave function, so π_{r1} has a maximum value.

As can be seen from Eq. 2, the Hessian matrix of π_{r2} is:

$$H_{c2} = \begin{bmatrix} \frac{\partial^2 \pi_{r2}}{\partial p_2^2} & \frac{\partial^2 \pi_{r2}}{\partial p_2 \partial b_2} \\ \frac{\partial^2 \pi_{r2}}{\partial b_2 \partial p_2} & \frac{\partial^2 \pi_{r2}}{\partial b_2^2} \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$$

$\Delta_1 = -2 < 0$; $\Delta_2 = 3 > 0$. Retailer 2's profit function π_{r2} is its decision variable p_2 with b_2 strict concave function, so π_{r2} has a maximum value.

Bringing p_1^{ct} , b_1^{ct} , p_2^{ct} , b_2^{ct} in Table 1 to Eq. 3, the Hessian matrix of π_m is:

$$H_{c3} = \begin{bmatrix} \frac{\partial^2 \pi_m}{\partial w_m^2} & \frac{\partial^2 \pi_m}{\partial w_m \partial \eta} & \frac{\partial^2 \pi_m}{\partial w_m \partial x_b} \\ \frac{\partial^2 \pi_m}{\partial \eta \partial w_m} & \frac{\partial^2 \pi_m}{\partial \eta^2} & \frac{\partial^2 \pi_m}{\partial \eta \partial x_b} \\ \frac{\partial^2 \pi_m}{\partial x_b \partial w_m} & \frac{\partial^2 \pi_m}{\partial x_b \partial \eta} & \frac{\partial^2 \pi_m}{\partial x_b^2} \end{bmatrix} = \begin{bmatrix} \frac{-8\beta+8}{2\beta-3} & \frac{2\beta-4}{2\beta-3} & \frac{-2}{2\beta-3} \\ \frac{2\beta-4}{2\beta-3} & \frac{-4\beta+8}{2\beta-3} & \frac{-2\beta+4}{2\beta-3} \\ \frac{-2}{2\beta-3} & \frac{-2\beta+4}{2\beta-3} & \frac{-8\beta+12}{2\beta-3} \end{bmatrix}$$

$\Delta_1 = \frac{-8\beta+8}{2\beta-3}$; $\Delta_2 = \frac{4(\beta-2)(7\beta-6)}{(2\beta-3)^2}$; $\Delta_3 = -\frac{16(\beta-2)(6\beta-5)}{(2\beta-3)^2}$. When $\Delta_1 < 0$, $\Delta_2 > 0$, $\Delta_3 > 0$, the manufacturer's profit function π_m is its decision variable w_m, η, x_b strict concave function, only having a maximum value. Making $\Delta_1 < 0$, $\Delta_2 > 0$, $\Delta_3 > 0$, the solution is obtained $\beta < \frac{5}{6}$, so we assume $\beta < \frac{5}{6}$, and the results are shown in Table 1.

When duopoly retailers play the Cournot game, the manufacturer sets a wholesale price threshold w'_m in order to guarantee its profit. Only when $w_m > w'_m$, the manufacturer gives the retailers a CSI bonus, otherwise it will not. As can be seen from Table 1 that when $\beta = 0.25$, it is a threshold of the wholesale price of the product.

$$w'_m = \frac{5(\alpha_1 + \alpha_2)}{14} \tag{4}$$

Table 1 Equilibrium under the Cournot of duopoly retailers (ct means Cournot)

Stage 1	wholesale price	$w_m^{ct} = -\frac{5(\alpha_1 + \alpha_2)}{4(6\beta - 5)}$
	assistance per unit CS effort	$x_b^{ct} = -\frac{\alpha_1 + \alpha_2}{4(6\beta - 5)}$
	bonus per unit CS effort	$\eta^{ct} = \begin{cases} \frac{(4\beta - 1)(\alpha_1 + \alpha_2)}{4(6\beta - 5)(\beta - 2)} & \beta > \frac{1}{4} \\ 0 & \beta \leq \frac{1}{4} \end{cases}$
Stage 2	retailer 1's selling price	$p_1^{ct} = \frac{A(w_m^{ct} + x_b^{ct} + \eta^{ct}) + 2B}{X}$
	retailer 1's CS effort	$b_1^{ct} = \frac{Dw_m^{ct} + E(x_b^{ct} + \eta^{ct}) + B}{X}$
	retailer 2's selling price	$p_2^{ct} = \frac{A(w_m^{ct} + x_b^{ct} + \eta^{ct}) + 2C}{X}$
	retailer 2's CS effort	$b_2^{ct} = \frac{Dw_m^{ct} + E(x_b^{ct} + \eta^{ct}) + C}{X}$

Note: $0 \leq \beta < \frac{5}{6}$, $X = 4\beta^2 - 9$, $A = -2\beta - 3$, $B = -2\beta\alpha_2 - 3\alpha_1$, $C = -2\beta\alpha_1 - 3\alpha_2$, $D = -2\beta^2 - \beta + 3$, $E = 2\beta^2 - \beta - 6$.

Proposition 1:

- (1) In the CS incentive plan, if $\beta > 0.25$, the manufacturer provides both CSI bonuses and CS assistance; If $\beta \leq 0.25$, the manufacturer only provides CS assistance.
- (2) When the degree of substitutability β between retailers increase, the manufacturer should increase CSI bonuses and CS assistance in the CS incentive plan.

Prove. In order to prove Proposition 1 (2), solving the partial derivative of x_b^{ct} and η^{ct} with respect to β in Table 1 is needed.

$$\frac{\partial x_b^{ct}}{\partial \beta} = \frac{24(\alpha_1 + \alpha_2)}{(24\beta - 20)^2} > 0, \quad \frac{\partial \eta^{ct}}{\partial \beta} = -\frac{(24\beta^2 - 12\beta - 13)(\alpha_1 + \alpha_2)}{4(6\beta - 5)^2(\beta - 2)^2} > 0 \tag{5}$$

Due to $\frac{\partial x_b^{ct}}{\partial \beta} > 0$ and $\frac{\partial \eta^{ct}}{\partial \beta} > 0$, x_b and η under Cournot game are increased with β , respectively.

Proposition 1 (1) indicates that when $\beta > 0.25$, the manufacturer should provide CSI bonuses and CS assistance to incentive retailers to improve CS, when $\beta \leq 0.25$, the manufacturer stops to use CSI bonuses. The larger β leads higher wholesale price of the manufacturers. Therefore, manufacturers are willing to provide CSI bonuses to retailers when wholesale prices are high. Conversely, when the wholesale price is low, the manufacturer will not provide the CSI bonus to the retailers. As can be seen from Table 1, $\beta = 0.25$ is a threshold for the wholesale price of the product. In addition, retailers alone bear the cost of the CS effort, and the benefits of high CS are shared by both parties. Therefore, manufacturers prefer to provide CS assistance to retailers to reduce the cost of the CS effort, and motivate retailers to make higher CS effort.

Proposition 1 (2) indicates that as β increase, the manufacturer should give retailers more CSI bonuses and CS assistance. Higher substitutability (β) between the two retailers is always preferred by the manufacturer, because the larger the β , the more similar the two retailers, and the easier it is for the manufacturer to manage the two retailers.

3.2 Retailers play the collusion game

In this section, it is assumed that duopoly retailers can collusion. Therefore, the total profit of retailers is:

$$\pi_R = \pi_{r1} + \pi_{r2} = (p_1 - w_m)D_1 - b_1(b_1 - x_b) + \eta b_1 + (p_2 - w_m)D_2 - b_2(b_2 - x_b) + \eta b_2 \tag{6}$$

Game sequence is:

- (1) The manufacturer determines w_m, η, x_b ;
- (2) Retailer 1 and retailer 2 determine p_i, b_i together, according to the given w_m, η, x_b .

In order to find the equilibrium of the subgame of the two-stage game, the backward induction method is adopted. The solution process is the same as 3.1. computes the Hessian matrix and obtain $\beta < \frac{5}{7}$, so we assume $\beta < \frac{5}{7}$, and the results are shown in Table 2.

Table 2 Equilibrium under the collusion of duopoly retailers (cn means collusion)

Stage 1	wholesale price	$w_m^{cn} = -\frac{\alpha_1 + \alpha_2}{4(\beta - 1)}$
	assistance per unit CS effort	$x_b^{cn} = -\frac{\alpha_1 + \alpha_2}{4(7\beta - 5)}$
	bonus per unit CS effort	$\eta^{cn} = 0$
Stage 2	retailer 1's selling price	$p_1^{cn} = \frac{Iw_m^{cn} - J(x_b^{cn} + \eta^{cn}) - 2G}{Y}$
	retailer 1's CS effort	$b_1^{cn} = \frac{Fw_m^{cn} - 2F(x_b^{cn} + \eta^{cn}) - G}{Y}$
	retailer 2's selling price	$p_2^{cn} = \frac{Iw_m^{cn} - J(x_b^{cn} + \eta^{cn}) - 2H}{Y}$
	retailer 2's CS effort	$b_2^{cn} = \frac{Fw_m^{cn} - 2F(x_b^{cn} + \eta^{cn}) - H}{Y}$

Note: $0 \leq \beta < \frac{5}{7}$. $Y = 16\beta^2 - 9$, $F = -4\beta^2 + \beta + 3$, $G = 4\beta\alpha_2 + 3\alpha_1$, $H = 4\beta\alpha_1 + 3\alpha_2$, $I = 8\beta^2 - 2\beta - 3$, $J = 4\beta + 3$.

Proposition 2:

- (1) The manufacturer only provides CS assistance in the CS incentive plan.
- (2) When the degree of substitutability β between retailers increase, the manufacturer should increase CS assistance in the CS incentive plan.

Prove. In order to prove Proposition 2 (2), solving the partial derivative of x_b^{cn} with respect to β in Table 2 is needed.

$$\frac{\partial x_b^{cn}}{\partial \beta} = \frac{28(\alpha_1 + \alpha_2)}{(28\beta - 20)^2} > 0 \tag{7}$$

Due to $\frac{\partial x_b^{cn}}{\partial \beta} > 0$, x_b under collusion game is increased with β .

Proposition 2 (1) indicates that when retailers adopt a collusion approach, manufacturers only provide CS assistance to retailers. The results of Huang *et al.* [11] showed that retailers' collusion behaviour will increase their selling prices, thus causing the manufacturer to encounter lower profits, and the retailer's collusion behaviour will cause the manufacturer to charge a lower wholesale price. Therefore, the manufacturer will only provide CS assistance to retailers and cancel the CSI bonus.

Proposition 2 (2) further indicates that when β increase, the manufacturer should give retailers more CSI bonuses and CS assistance.

3.3 Impact of the strategy

This section discusses the impact of the Cournot and collusion behaviour of duopoly retailers on each of the equilibrium solutions in Table 1 and Table 2, and analyse the preferences of manufacturers and retailers. The following discussion is in $\beta < \frac{5}{7}$ the case was carried out, in which the superscript *ct* represents Cournot, and the superscript *cn* represents collusion.

Proposition 3: $w_m^{cn} < w_m^{ct}$.

Prove. w_m^{cn} with w_m^{ct} make a difference

$$w_m^{cn} - w_m^{ct} = - \frac{\beta(\alpha_1 + \alpha_2)}{24\beta^2 - 44\beta + 20} < 0 \tag{8}$$

The difference between the two is less than 0, so the proposition is proved.

Proposition 3 indicates that the retailer's collusion behaviour will cause the manufacturer to charge lower wholesale price. Consistent with the conclusions in literature [11].

Proposition 4: $\pi_m^{cn} < \pi_m^{ct}$.

Prove. π_m^{cn} with π_m^{ct} make a difference

$$\pi_m^{cn} - \pi_m^{ct} = \begin{cases} - \frac{(588\beta^4 - 3896\beta^3 + 6712\beta^2 - 4484\beta + 1045)(\alpha_1 + \alpha_2)^2}{8(6\beta^2 - 17\beta + 10)(4\beta - 3)(7\beta - 5)^2} < 0 & \beta > \frac{1}{4} \\ - \frac{(7056\beta^5 - 45968\beta^4 + 98116\beta^3 - 95380\beta^2 + 43905\beta - 7800)(\alpha_1 + \alpha_2)^2}{8(2\beta - 3)(4\beta - 3)(7\beta - 5)^2(6\beta - 5)^2} < 0 & \beta \leq \frac{1}{4} \end{cases} \tag{9}$$

The difference between the two is less than 0, so the proposition is proved.

Proposition 4 indicates that the retailer's collusion behaviour will cause the manufacturer to suffer lower profits, because the collusion behaviour of retailers will increase selling prices and reduce sales.

Proposition 5: $p_i^{cn} > p_i^{ct}$, if $\beta > 0.25$; $p_1^{cn} > p_1^{ct}$, if $\beta \leq 0.25$ and $\alpha_1 \leq \alpha_2$; $p_2^{cn} > p_2^{ct}$, if $\beta \leq 0.25$ and $\alpha_1 \geq \alpha_2$; $b_i^{cn} > b_i^{ct}$.

Prove. First p_1^{cn} with p_1^{ct} make a difference

$$p_1^{cn} - p_1^{ct} = \begin{cases} \frac{\alpha_1 + \alpha_2}{2(4\beta - 3)} - \frac{7\alpha_1 + 7\alpha_2}{4(7\beta - 5)} - \frac{\alpha_1 + \alpha_2}{4(\beta - 1)} + \frac{-\alpha_1 + \alpha_2}{2\beta + 3} + \frac{3\alpha_1 + 3\alpha_2}{2(6\beta - 5)} + \frac{\alpha_1 + \alpha_2}{4(\beta - 2)} + \frac{\alpha_1 - \alpha_2}{4\beta + 3} > 0 & \beta > \frac{1}{4} \\ \frac{\alpha_1 + \alpha_2}{2(4\beta - 3)} - \frac{7\alpha_1 + 7\alpha_2}{4(7\beta - 5)} - \frac{\alpha_1 + \alpha_2}{4(\beta - 1)} + \frac{-\alpha_1 + \alpha_2}{2\beta + 3} + \frac{9\alpha_1 + 9\alpha_2}{8(6\beta - 5)} + \frac{5\alpha_1 + 5\alpha_2}{8(2\beta - 3)} + \frac{\alpha_1 - \alpha_2}{4\beta + 3} & \beta \leq \frac{1}{4} \end{cases} \tag{10}$$

When $\beta \leq \frac{1}{4}$, $\frac{\alpha_1+\alpha_2}{2(4\beta-3)} - \frac{7\alpha_1+7\alpha_2}{4(7\beta-5)} - \frac{\alpha_1+\alpha_2}{4(\beta-1)} + \frac{9\alpha_1+9\alpha_2}{8(6\beta-5)} + \frac{5\alpha_1+5\alpha_2}{8(2\beta-3)} > 0$; So if $\frac{-\alpha_1+\alpha_2}{2\beta+3} + \frac{\alpha_1-\alpha_2}{4\beta+3} = -\frac{2(\alpha_1-\alpha_2)\beta}{8\beta^2+18\beta+9} \geq 0$, $p_1^{cn} - p_1^{ct}$ is greater than 0. If $\frac{-\alpha_1+\alpha_2}{2\beta+3} + \frac{\alpha_1-\alpha_2}{4\beta+3} \geq 0$, then $\alpha_1 \leq \alpha_2$.

Similarly, p_2^{cn} with p_2^{ct} make a difference

$$p_2^{cn} - p_2^{ct} = \begin{cases} \frac{\alpha_1+\alpha_2}{2(4\beta-3)} - \frac{7\alpha_1+7\alpha_2}{4(7\beta-5)} - \frac{\alpha_1+\alpha_2}{4(\beta-1)} + \frac{\alpha_1-\alpha_2}{2\beta+3} + \frac{3\alpha_1+3\alpha_2}{2(6\beta-5)} + \frac{\alpha_1+\alpha_2}{4(\beta-2)} + \frac{-\alpha_1+\alpha_2}{4\beta+3} > 0 & \beta > \frac{1}{4} \\ \frac{\alpha_1+\alpha_2}{2(4\beta-3)} - \frac{7\alpha_1+7\alpha_2}{4(7\beta-5)} - \frac{\alpha_1+\alpha_2}{4(\beta-1)} + \frac{\alpha_1-\alpha_2}{2\beta+3} + \frac{9\alpha_1+9\alpha_2}{8(6\beta-5)} + \frac{5\alpha_1+5\alpha_2}{8(2\beta-3)} + \frac{-\alpha_1+\alpha_2}{4\beta+3} & \beta \leq \frac{1}{4} \end{cases} \quad (11)$$

When $\beta \leq \frac{1}{4}$, $\frac{\alpha_1+\alpha_2}{2(4\beta-3)} - \frac{7\alpha_1+7\alpha_2}{4(7\beta-5)} - \frac{\alpha_1+\alpha_2}{4(\beta-1)} + \frac{9\alpha_1+9\alpha_2}{8(6\beta-5)} + \frac{5\alpha_1+5\alpha_2}{8(2\beta-3)} > 0$; So if $\frac{\alpha_1-\alpha_2}{2\beta+3} + \frac{-\alpha_1+\alpha_2}{4\beta+3} = \frac{2(\alpha_1-\alpha_2)\beta}{8\beta^2+18\beta+9} \geq 0$, $p_2^{cn} - p_2^{ct}$ is greater than 0. If $\frac{\alpha_1-\alpha_2}{2\beta+3} + \frac{-\alpha_1+\alpha_2}{4\beta+3} \geq 0$, then $\alpha_1 \geq \alpha_2$.

At last, subtract b_1^{cn} from b_1^{ct} , and b_2^{cn} from b_2^{ct} separately.

$$b_1^{cn} - b_1^{ct} = \begin{cases} \frac{\alpha_1+\alpha_2}{4(4\beta-3)} - \frac{\alpha_1+\alpha_2}{7\beta-5} + \frac{-\alpha_1+\alpha_2}{2(2\beta+3)} + \frac{\alpha_1+\alpha_2}{2(6\beta-5)} + \frac{\alpha_1-\alpha_2}{2(4\beta+3)} > 0 & \beta > \frac{1}{4} \\ \frac{5\alpha_1+5\alpha_2}{16(2\beta-3)} + \frac{\alpha_1+\alpha_2}{4(4\beta-3)} - \frac{\alpha_1+\alpha_2}{7\beta-5} + \frac{-\alpha_1+\alpha_2}{2(2\beta+3)} + \frac{\alpha_1+\alpha_2}{16(6\beta-5)} + \frac{\alpha_1-\alpha_2}{2(4\beta+3)} > 0 & \beta \leq \frac{1}{4} \end{cases} \quad (12)$$

$$b_2^{cn} - b_2^{ct} = \begin{cases} \frac{\alpha_1+\alpha_2}{4(4\beta-3)} - \frac{\alpha_1+\alpha_2}{7\beta-5} + \frac{\alpha_1-\alpha_2}{2(2\beta+3)} + \frac{\alpha_1+\alpha_2}{2(6\beta-5)} + \frac{-\alpha_1+\alpha_2}{2(4\beta+3)} > 0 & \beta > \frac{1}{4} \\ \frac{5\alpha_1+5\alpha_2}{16(2\beta-3)} + \frac{\alpha_1+\alpha_2}{4(4\beta-3)} - \frac{\alpha_1+\alpha_2}{7\beta-5} + \frac{\alpha_1-\alpha_2}{2(2\beta+3)} + \frac{\alpha_1+\alpha_2}{16(6\beta-5)} + \frac{-\alpha_1+\alpha_2}{2(4\beta+3)} > 0 & \beta \leq \frac{1}{4} \end{cases} \quad (13)$$

The difference between the above two is greater than 0, so the proposition is proved.

Proposition 5 indicates that the retailer's collusion behaviour will cause them to charge higher selling prices and make greater CS efforts. When retailers raise their selling prices, sales will decrease. At this time, in order to avoid excessive sales reduction, retailers will make greater effort to increase CS to increase sales. In other words, when retailers choose collusion behaviour, they will confrontation the reduction in sales due to rising selling prices by increasing CS efforts.

Proposition 6: $\pi_{ri}^{cn} > \pi_{ri}^{ct}$.

Prove. Subtract π_{r1}^{cn} from π_{r1}^{ct} , and π_{r2}^{cn} from π_{r2}^{ct} separately.

$$\pi_{r1}^{cn} - \pi_{r1}^{ct} = \begin{cases} \frac{-3(\alpha_1-\alpha_2)^2}{4(2\beta+3)^2} + \frac{-3(\alpha_1+\alpha_2)^2}{16(6\beta-5)^2} + \frac{-(\alpha_1+\alpha_2)^2}{16(\beta-2)^2} + \frac{(\alpha_1+\alpha_2)^2}{8(7\beta-5)^2} - \frac{21(\alpha_1^2-\alpha_2^2)(84\beta^3-399\beta^2+275\beta+5)}{328(2\beta+3)(7\beta-5)(6\beta-5)(\beta-2)} + \frac{125\alpha_1^2-164\alpha_1\alpha_2+39\alpha_2^2}{328(4\beta+3)} + \frac{\alpha_1^2-2\alpha_1\alpha_2+3\alpha_2^2}{16(4\beta-3)} > 0 & \beta > \frac{1}{4} \\ \frac{-3(\alpha_1-\alpha_2)^2}{4(2\beta+3)^2} + \frac{-7(\alpha_1+\alpha_2)^2}{256(6\beta-5)^2} + \frac{-75(\alpha_1+\alpha_2)^2}{256(2\beta-3)^2} + \frac{(\alpha_1+\alpha_2)^2}{8(7\beta-5)^2} + \frac{\alpha_1^2-\alpha_2^2}{7(2\beta+3)} + \frac{95\alpha_1^2+30\alpha_1\alpha_2-65\alpha_2^2}{512(2\beta-3)} + \frac{125\alpha_1^2-164\alpha_1\alpha_2+39\alpha_2^2}{328(4\beta+3)} + \frac{147(\alpha_1^2-\alpha_2^2)}{328(7\beta-5)} + \frac{\alpha_1^2-2\alpha_1\alpha_2+3\alpha_2^2}{16(4\beta-3)} - \frac{459\alpha_1^2+630\alpha_1\alpha_2+171\alpha_2^2}{3584(6\beta-5)} > 0 & \beta \leq \frac{1}{4} \end{cases} \quad (14)$$

$$\pi_{r2}^{cn} - \pi_{r2}^{ct} = \begin{cases} \frac{-3(\alpha_1-\alpha_2)^2}{4(2\beta+3)^2} + \frac{-3(\alpha_1+\alpha_2)^2}{16(6\beta-5)^2} + \frac{-(\alpha_1+\alpha_2)^2}{16(\beta-2)^2} + \frac{(\alpha_1+\alpha_2)^2}{8(7\beta-5)^2} - \frac{21(\alpha_1^2-\alpha_2^2)(84\beta^3-399\beta^2+275\beta+5)}{328(2\beta+3)(7\beta-5)(6\beta-5)(\beta-2)} + \frac{39\alpha_1^2-164\alpha_1\alpha_2+125\alpha_2^2}{328(4\beta+3)} + \frac{3\alpha_1^2-2\alpha_1\alpha_2+\alpha_2^2}{16(4\beta-3)} > 0 & \beta > \frac{1}{4} \\ \frac{-3(\alpha_1-\alpha_2)^2}{4(2\beta+3)^2} + \frac{-7(\alpha_1+\alpha_2)^2}{256(6\beta-5)^2} + \frac{-75(\alpha_1+\alpha_2)^2}{256(2\beta-3)^2} + \frac{(\alpha_1+\alpha_2)^2}{8(7\beta-5)^2} + \frac{\alpha_1^2-\alpha_2^2}{7(2\beta+3)} - \frac{171\alpha_1^2+630\alpha_1\alpha_2+459\alpha_2^2}{3584(6\beta-5)} + \frac{-3\alpha_1^2-2\alpha_1\alpha_2+\alpha_2^2}{16(4\beta-3)} + \frac{-65\alpha_1^2+30\alpha_1\alpha_2-95\alpha_2^2}{512(2\beta-3)} + \frac{147(\alpha_1^2-\alpha_2^2)}{328(7\beta-5)} + \frac{39\alpha_1^2-164\alpha_1\alpha_2+125\alpha_2^2}{328(4\beta+3)} > 0 & \beta \leq \frac{1}{4} \end{cases} \quad (15)$$

The difference between the above two is greater than 0, so the proposition is proved.

Proposition 6 indicates that regardless of the initial reputation of the two retailers and the degree of substitutability between the two retailers, they will choose collusion behaviour because collusion will result in higher profits.

Besides discussing the respective profits of manufacturers and retailers, the profits of the entire supply chain should also be taken into account. The total profit of the supply chain π_A^j is:

$$\pi_A^j = \pi_m^j + \pi_{r1}^j + \pi_{r2}^j \quad (j = ct \text{ or } cn) \quad (16)$$

Proposition 7: $\pi_A^{cn} > \pi_A^{ct}$, If $\beta \geq 0.20$; $\pi_A^{cn} < \pi_A^{ct}$, If $\beta \leq 0.15$.

Prove. π_A^{cn} and π_A^{ct} make a difference

$$\pi_A^{cn} - \pi_A^{ct} = \begin{cases} \frac{(\alpha_1 + \alpha_2)^2}{8(4\beta - 3)} + \frac{(\alpha_1 - \alpha_2)^2}{2(4\beta + 3)} + \frac{(\alpha_1 + \alpha_2)^2}{8(7\beta - 5)^2} - \frac{7(\alpha_1 + \alpha_2)^2}{8(7\beta - 5)} - \frac{3(\alpha_1 - \alpha_2)^2}{2(2\beta + 3)^2} - \frac{3(\alpha_1 + \alpha_2)^2}{8(6\beta - 5)^2} + \frac{3(\alpha_1 + \alpha_2)^2}{4(6\beta - 5)} - \frac{(\alpha_1 + \alpha_2)^2}{8(\beta - 2)} - \frac{(\alpha_1 + \alpha_2)^2}{8(\beta - 2)^2} > 0 & \beta > \frac{1}{4} \\ \frac{(\alpha_1 + \alpha_2)^2}{8(4\beta - 3)} + \frac{(\alpha_1 - \alpha_2)^2}{2(4\beta + 3)} + \frac{(\alpha_1 + \alpha_2)^2}{8(7\beta - 5)^2} - \frac{7(\alpha_1 + \alpha_2)^2}{8(7\beta - 5)} - \frac{3(\alpha_1 - \alpha_2)^2}{2(2\beta + 3)^2} + \frac{49(\alpha_1 + \alpha_2)^2}{128(6\beta - 5)^2} + \frac{255(\alpha_1 + \alpha_2)^2}{256(6\beta - 5)} - \frac{85(\alpha_1 + \alpha_2)^2}{256(2\beta - 3)} - \frac{75(\alpha_1 + \alpha_2)^2}{128(2\beta - 3)^2} & \beta \leq \frac{1}{4} \end{cases} \quad (17)$$

When $\beta > \frac{1}{4}$, $\pi_A^{cn} - \pi_A^{ct} > 0$; But when $\beta \leq \frac{1}{4}$, $\pi_A^{cn} - \pi_A^{ct}$ there are both greater than 0 and less than 0, so discussions were made at $\beta=0.05, 0.1, 0.15, 0.20, 0.25$. When $\beta=0.05, 0.1, 0.15$ $\pi_A^{cn} - \pi_A^{ct} < 0$; When $\beta=0.20, 0.25$ $\pi_A^{cn} - \pi_A^{ct} > 0$, so the proposition is proved.

Proposition 7 indicates that when β is large, collusion will make the entire supply chain getting more profit. When β is small, Cournot will make the entire supply chain getting more profit. Because there is an alternative relationship between the needs of the two retailers. When this alternative relationship is stronger (β larger), it will have a greater impact on the respective needs of the two retailers. The collusion between the two parties will reduce the impact on demand, thus making the entire supply chain more profits. When the substitution relationship is weaker (β smaller), the impact on the respective needs of the two retailers is smaller. Therefore, the two parties are unwilling to collusion, they all hope to maximize their profits through competition, thus making the entire supply chain more profits.

3.4 The impact of substitutability β on decision results

This section discusses the impact of substitutability β on optimal wholesale price, selling price, profit, CS effort, and CS incentives. Let $\alpha_1 = 100, \alpha_2 = 90$, the calculation results are shown in the figures Figs. 1 to 6. Fig. 1 and Fig. 2 show that the wholesale price and profit of the manufacturer under two games increase with the increase of β . When $\beta = 0$, the wholesale price and profit under the two games are the same. With the increase of β , the manufacturer's wholesale price difference between the two games will also increase, while the manufacturer's profit difference between the two games will increase first and then decrease. Therefore, the cross-price effect of downstream retailers will have a significant impact on the manufacturer's wholesale price and profit. When $\beta > 0$, the wholesale price and profit of the manufacturer under the Cournot game of downstream retailers are larger, which is consistent with Proposition 3 and Proposition 4.

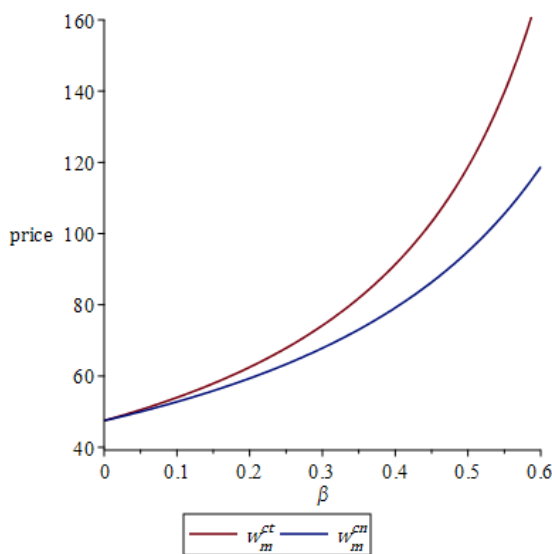


Fig. 1 The optimal wholesale price under two games

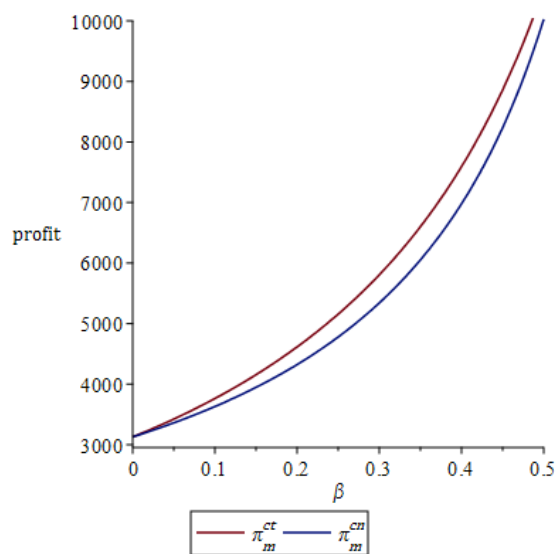


Fig. 2 The optimal profit of the manufacturer under two games

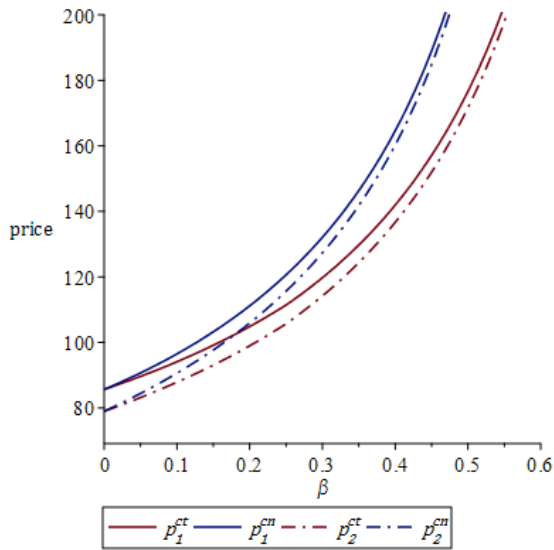


Fig. 3 The optimal selling price under two games

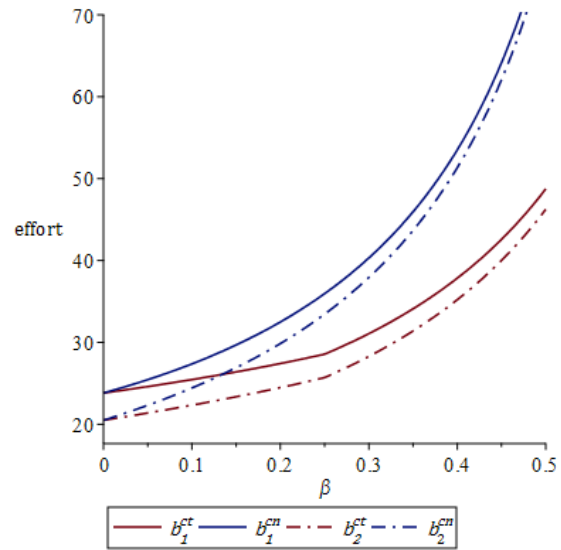


Fig. 4 The optimal CS effort under two games

Fig. 3, Fig. 4 and Fig. 5 show that the optimal selling price, CS effort and profit of retailer i under two games increase with the increase of β . When $\beta = 0$, the selling price, CS effort and profit of the same retailer under the two games are the same. With the increase of β , the difference of selling price, CS effort and profit of each retailer between the two games will also increase. When $\beta > 0$, the selling price, CS effort and profit of the retailer i under the collusion game of downstream retailers are larger. This is because when downstream retailers play collusion game, they will collude to increase their own interests and squeeze the interests of the manufacturer. This is why the collusion of downstream retailers is illegal in some countries.

We can see from Fig. 6 that the manufacturer's CS assistance per unit will increase with the increase of β under the two games. When $\beta = 0$, the two are equal, when $\beta > 0$, the CS assistance per unit is larger than in the collusion game of retailers. When $0 < \beta \leq 0.25$, the manufacturer's CSI bonus per unit was 0. When $\beta > 0.25$, the CSI bonus with the Cournot game will still be 0, while the CSI bonus in collusion game will increase with the increase of β , and the increase rate is faster, which will gradually be greater than the CS assistance in the Cournot game.

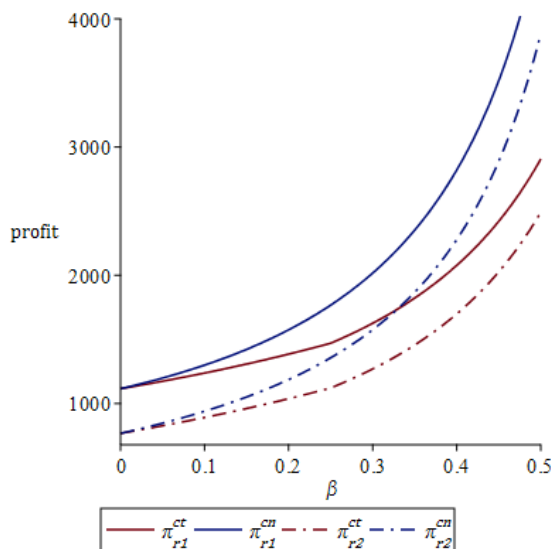


Fig. 5 The optimal profit of retailers under two games

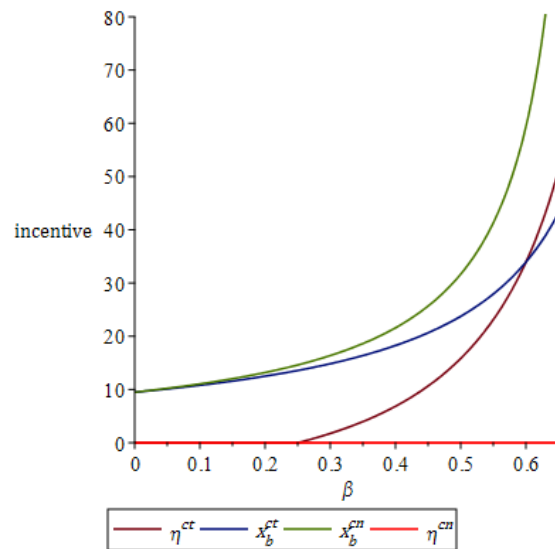


Fig. 6 The optimal CS incentive under two games

4. The impact of manufacturer CS incentives

The CS incentive is provided by the manufacturer, therefore, from the perspective of the manufacturer, we study the impact of the CS incentive on the manufacturer. First, we must find the manufacturer's decision when the CS incentive is not provided, and then compare them with the decision in section 3. Here, we use (–) to represent the results in the case of no CS incentives.

Table 3 Equilibrium under duopoly retailers' two games without CS incentive

Stage	Optimal value	Cournot ($j = ct$)	Collusion ($j = cn$)
Stage 1	\overline{w}_m^j	$-\frac{\alpha_1 + \alpha_2}{4(\beta - 1)}$	$-\frac{\alpha_1 + \alpha_2}{4(\beta - 1)}$
Stage 2	\overline{p}_1^j	$-\frac{(2\beta + 3)\overline{w}_m^j + 4\beta\alpha_2 + 6\alpha_1}{4\beta^2 - 9}$	$\frac{(8\beta^2 + 2\beta - 3)\overline{w}_m^j - 8\beta\alpha_2 - 6\alpha_1}{4\beta^2 - 9}$
	\overline{b}_1^j	$-\frac{(2\beta^2 + \beta - 3)\overline{w}_m^j + 2\beta\alpha_2 + 3\alpha_1}{4\beta^2 - 9}$	$-\frac{(4\beta^2 - \beta - 3)\overline{w}_m^j + 4\beta\alpha_2 + 3\alpha_1}{4\beta^2 - 9}$
	\overline{p}_2^j	$-\frac{(2\beta + 3)\overline{w}_m^j + 4\beta\alpha_1 + 6\alpha_2}{4\beta^2 - 9}$	$\frac{(8\beta^2 + 2\beta - 3)\overline{w}_m^j - 8\beta\alpha_1 - 6\alpha_2}{4\beta^2 - 9}$
	\overline{b}_2^j	$-\frac{(2\beta^2 + \beta - 3)\overline{w}_m^j + 2\beta\alpha_1 + 3\alpha_2}{4\beta^2 - 9}$	$-\frac{(4\beta^2 - \beta - 3)\overline{w}_m^j + 4\beta\alpha_1 + 3\alpha_2}{4\beta^2 - 9}$

4.1 Game analysis when manufacturers do not provide CS incentives

Assuming that the manufacturer will not provide CS incentives to retailers, let's analyse this two-stage decision-making problem, in which the demand function is same as Eq. 1. Retailer i the profit function $\overline{\pi}_{ri}$ is:

$$\overline{\pi}_{ri} = (p_i - w_m)D_i - b_i^2 \tag{18}$$

The manufacturer 's profit function $\overline{\pi}_m$ is:

$$\overline{\pi}_m = w_m(D_1 + D_2) \tag{19}$$

Similarly, we use the backward induction method to obtain the equilibrium without any CS incentive plan, and the results are shown in Table 3.

Bringing the equilibrium in Table 3 into Eq. 19, the manufacturer's profits in the two models without any CS incentive plan respectively are:

$$\overline{\pi}_m^{ct} = \frac{(\alpha_1 + \alpha_2)^2}{8\beta^2 - 20\beta + 12} \tag{20}$$

$$\overline{\pi}_m^{cn} = -\frac{(\alpha_1 + \alpha_2)^2}{16\beta - 12} \tag{21}$$

4.2 The impact of CS incentives on manufacturer's profit and wholesale price

We compare the wholesale price and profit of the manufacturer in the two cases with or without CS incentives.

Proposition 8: The manufacturer's CS incentives may or may not increase its wholesale price, it depends on which game the duopoly retailer adopts, but it will certainly obtain higher profits.

$$w_m^{ct} - \overline{w}_m^{ct} = \frac{(\alpha_1 + \alpha_2)\beta}{24\beta^2 - 44\beta + 20} > 0 \tag{22}$$

$$w_m^{cn} - \overline{w}_m^{cn} = 0 \tag{23}$$

$$\frac{\overline{\pi}_m^{ct} - \overline{\pi}_m^{cn}}{\overline{\pi}_m^{ct}} = \begin{cases} \frac{2\beta^2 - \beta + 1}{12\beta^2 - 34\beta + 20} > 0 & \beta > \frac{1}{4} \\ \frac{-2\beta^2 + 1}{(6\beta - 5)^2} > 0 & \beta \leq \frac{1}{4} \end{cases} \tag{24}$$

$$\frac{\overline{\pi}_m^{cn} - \overline{\pi}_m^{cn}}{\overline{\pi}_m^{cn}} = \frac{3 - 2\beta}{2(7\beta - 5)^2} > 0 \tag{25}$$

From Eq. 22 and Eq. 23, we can see that when the retailers adopt a Cournot game, the manu-

facturer's CS incentive plan will increase its wholesale price; when the retailers adopt a collusion game, its wholesale price remains unchanged. Because Eq. 24 is greater than 0, $\overline{\pi}_m^{ct} > 0$, so $\pi_m^{ct} - \overline{\pi}_m^{ct} > 0$; Because Eq. 25 is greater than 0, $\overline{\pi}_m^{cn} > 0$, so $\pi_m^{cn} - \overline{\pi}_m^{cn} > 0$. Therefore, no matter which game the retailers adopt, the manufacturer's CS incentive plan will make it obtain higher profits.

4.3 The impact of CS incentives on consumer's demand

Besides whole sale price, the demand is also an important issue for the future suces of the manufacturer. Therefore, in this section, we will discuss the impact of CS incentives on consumer's demand. The consumer's demand function is:

$$D_m = D_1 + D_2 = \alpha_1 + \alpha_2 - p_1 - p_2 + b_1 + b_2 + \beta(p_1 + p_2) \tag{26}$$

As shown in Fig. 7, no matter what game the two retailers adopt, the manufacturer's CS incentive plan will increase their demand, and its demand will increase with the increase of β . In the case that the manufacturer provides CS incentives, when $\beta = 0$, the demand in the two games is equal. With the increase of β , first the demand in the Cournot game is larger, and then the demand in the collusion game went up quickly and exceed that in the Cournot game. In the case that the manufacturer does not provide CS incentives, when $\beta = 0$, the demand in the two games is equal. Similarly, with the increase of β , the demand in the Cournot game is larger at the beginning, but then the demand in the collusion game becomes larger than that in the Cournot game.

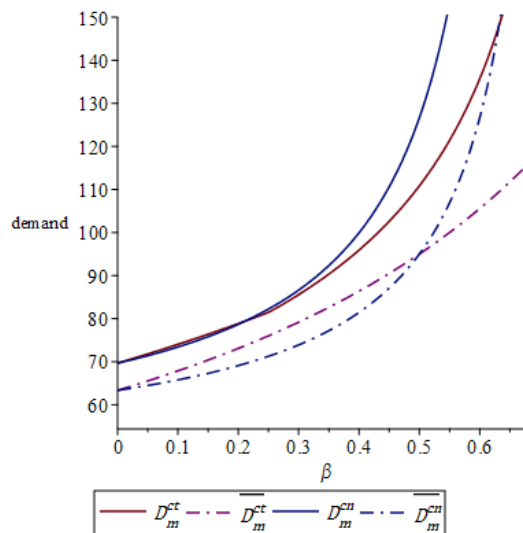


Fig. 7 Manufacturer's demand in four cases

5. Conclusion

This paper examined a two-echelon supply chain system consisting of a manufacturer and duopoly retailers. The manufacturer offers retailers two types of CS incentives to enable retailers to improve CS. First, the paper establishes two models to study the best CS incentive plan for duopoly retailers in the context of Cournot and collusion. Then, compare and analyse the equilibrium results of the two models. Finally, we study the impact of the CS incentive plan on the manufacturer. The results showed that:

- When the two retailers take collusion behaviour, the manufacturer only provides CS assistance to the retailers but no CSI bonus.
- In the Cournot game condition if the wholesale price is greater than the threshold, the manufacturer provides CSI bonus and CS assistance to the retailers; if the wholesale price is less than the threshold, the manufacturer only provides CS assistance to the retailers.
- In both models, when the degree of substitutability between retailers increases, the manufacturer should always increase CS assistance in the CS incentive plan. In the Cournot

model, when the degree of substitutability between retailers increases, the manufacturer should increase CSI bonuses in the CS incentive plan.

- Retailers always like collusion behaviour, because collusion behaviour will cause manufacturers to lower wholesale prices, which will enable retailers to increase selling prices and make greater CS efforts, thus enabling retailers to higher profit. Manufacturers always like Cournot behaviour, because retailers' Cournot behaviour will make them more profitable.
- The decision results of the best wholesale price, profit, selling price and CS effort will increase as the degree of substitution between retailers increases.
- The manufacturer's CS incentives may or may not increase its wholesale price, which mainly depends on what kind of game the retailers adopt, but it will certainly obtain more demand and higher profits.

According to the above analysis and proposition, the manufacturer should fully consider the marketing mix decisions between retailers when formulating CS incentive plans, and try to cooperate with retailers with a high degree of substitution. Although the model proposed in this paper is more in line with the enterprise production management practice, there are still some limitations. In the demand function, we assumed that the sensitivity of the CS effort is 1, but it may not always be the case in real life. Therefore, the demand function needs to be improved. The downstream market is generally composed of multiple retailers, and this study only considers two retailers. In addition, this paper only considers the Cournot and collusion game of duopoly retailers, and future research can consider Stackelberg game.

Acknowledgement

This work is supported by National Natural Science Foundation of China (No. 71704151), Research funding of Hebei Key Research Institute of Humanities and Social Sciences at Universities (JJ1907).

References

- [1] Wang, Y., Zhao, P. (2010). Relationship between customer satisfaction and shareholder value, *Journal of Management Sciences in China*, Vol. 13, No. 6, 54-63.
- [2] Strenitzerová, M., Gaña, J. (2018). Customer satisfaction and loyalty as a part of customer-based corporate sustainability in the sector of mobile communications services, *Sustainability*, Vol. 10, No. 5, Article No. 1657, [doi: 10.3390/su10051657](https://doi.org/10.3390/su10051657).
- [3] Tam, J.L.M. (2011). The moderating effects of purchase importance in customer satisfaction process: An empirical investigation, *Journal of Consumer Behaviour*, Vol. 10, No. 4, 205-215, [doi: 10.1002/cb.330](https://doi.org/10.1002/cb.330).
- [4] Fornell, C. (1992). A national customer satisfaction barometer: The Swedish experience, *Journal of Marketing*, Vol. 56, No. 1, 6-21, [doi: 10.1177/002224299205600103](https://doi.org/10.1177/002224299205600103).
- [5] Hu, H., Djebarni, R., Zhao, X., Xiao, L., Flynn, B. (2017). Effect of different food recall strategies on consumers' reaction to different recall norms: A comparative study, *Industrial Management & Data Systems*, Vol. 117, No. 9, 2045-2063, [doi: 10.1108/IMDS-10-2016-0464](https://doi.org/10.1108/IMDS-10-2016-0464).
- [6] Chu, W., Desai, P.S. (1995). Channel coordination mechanisms for customer satisfaction, *Marketing Science*, Vol. 14, No. 4, 343-359, [doi: 10.1287/mksc.14.4.343](https://doi.org/10.1287/mksc.14.4.343).
- [7] Wang, C.X., Qian, Z., Zhao, Y. (2018). Impact of manufacturer and retailer's market pricing power on customer satisfaction incentives in supply chains, *International Journal of Production Economics*, Vol. 205, 98-112, [doi: 10.1016/j.ijpe.2018.08.034](https://doi.org/10.1016/j.ijpe.2018.08.034).
- [8] Cardozo, R.N. (1965). An experimental study of customer effort, expectation, and satisfaction, *Journal of Marketing Research*, Vol. 2, No. 3, 244-249, [doi: 10.1177/002224376500200303](https://doi.org/10.1177/002224376500200303).
- [9] Hauser, J.R., Simester, D.I., Wernerfelt, B. (1994). Customer satisfaction incentives, *Marketing Science*, Vol. 13, No. 4, 327-350, [doi: 10.1287/mksc.13.4.327](https://doi.org/10.1287/mksc.13.4.327).
- [10] Yang, S.-L., Zhou, Y.-W. (2006). Two-echelon supply chain models: Considering duopolistic retailers' different competitive behaviors, *International Journal of Production Economics*, Vol. 103, No. 1, 104-116, [doi: 10.1016/j.ijpe.2005.06.001](https://doi.org/10.1016/j.ijpe.2005.06.001).
- [11] Huang, H., Ke, H., Wang, L. (2016). Equilibrium analysis of pricing competition and cooperation in supply chain with one common manufacturer and duopoly retailers, *International Journal of Production Economics*, Vol. 178, 12-21, [doi: 10.1016/j.ijpe.2016.04.022](https://doi.org/10.1016/j.ijpe.2016.04.022).
- [12] Modak, N.M., Panda, S., Sana, S.S. (2016). Two-echelon supply chain coordination among manufacturer and duopolies retailers with recycling facility, *The International Journal of Advanced Manufacturing Technology*, Vol. 87, No. 5-8, 1531-1546, [doi: 10.1007/s00170-015-8094-y](https://doi.org/10.1007/s00170-015-8094-y).

- [13] Herbon, A. (2018). Prices, profits and market shares of two competing retailers under consumer information asymmetry, *Computers & Industrial Engineering*, Vol. 124, 195-206, doi: [10.1016/j.cie.2018.07.010](https://doi.org/10.1016/j.cie.2018.07.010).
- [14] Li, Q.Z., Fan, X.W., Huang, W.J., Kwangseek, C. (2017). Collaborative supply model and case simulation in a two-level assemble-to-order system in the context of global purchasing, *International Journal of Simulation Modelling*, Vol. 16, No. 3, 471-483, doi: [10.2507/IJSIMM16\(3\)9.393](https://doi.org/10.2507/IJSIMM16(3)9.393).
- [15] Lang, X., Shao, X. (2019). Research on retailers' assortment decision under competition, *Systems Engineering-Theory & Practice*, Vol. 39, No. 5, 1105-1116, doi: [10.12011/1000-6788-2018-0155-12](https://doi.org/10.12011/1000-6788-2018-0155-12).
- [16] Chai, W., Sun, H., Wang, W., Wu, J. (2013). Price competition model in decentralized and centralized supply chains with demand disruption, *Journal of Industrial Engineering and Management*, Vol. 6, No. 1, 16-24, doi: [10.3926/jiem.616](https://doi.org/10.3926/jiem.616).
- [17] Saha, S., Panda, S., Modak, N.M., Basu, M. (2015). Mail-in-rebate coupled with revenue sharing and downward direct discount for supply chain coordination, *International Journal of Operational Research*, Vol. 23, No. 4, 451-476, doi: [10.1504/IJOR.2015.070145](https://doi.org/10.1504/IJOR.2015.070145).
- [18] Modak, N.M., Panda, S., Sana, S.S. (2016). Three-echelon supply chain coordination considering duopolistic retailers with perfect quality products, *International Journal of Production Economics*, Vol. 182, 564-578, doi: [10.1016/j.ijpe.2015.05.021](https://doi.org/10.1016/j.ijpe.2015.05.021).
- [19] Ingene, C.A., Parry, M.E. (1995). Channel coordination when retailers compete, *Marketing Science*, Vol. 14, No. 4, 360-377, doi: [10.1287/mksc.14.4.360](https://doi.org/10.1287/mksc.14.4.360).
- [20] Li, B., Zhu, M., Jiang, Y., Li, Z. (2016). Pricing policies of a competitive dual-channel green supply chain, *Journal of Cleaner Production*, Vol. 112, Part 3, 2029-2042, doi: [10.1016/j.jclepro.2015.05.017](https://doi.org/10.1016/j.jclepro.2015.05.017).
- [21] He, X., Li, W., Nie, K. (2013). Pricing and coordination research for TPL based on different logistics service level, *Journal of System and Management Sciences*, Vol. 3, No. 2, 35-44.
- [22] Pang, Q.H., Wu, X.Y., Tan, M.L., Cao, X.Y. (2015). Supply chain coordination using revenue-sharing contract with distributor's effort dependent demand, *International Journal of Simulation Modelling*, Vol. 14, No. 2, 335-348, doi: [10.2507/IJSIMM14\(2\)CO8](https://doi.org/10.2507/IJSIMM14(2)CO8).
- [23] Li, S.X., Huang, Z., Ashley, A. (1996). Improving buyer-seller system cooperation through inventory control, *International Journal of Production Economics*, Vol. 43, No. 1, 37-46, doi: [10.1016/0925-5273\(95\)00183-2](https://doi.org/10.1016/0925-5273(95)00183-2).
- [24] Wang, X.L., Hu, S.Q., Liu, X.B. (2019). Green supply chain game models and contract coordination with government incentives, *Industrial Engineering Journal*, Vol. 22, No. 6, 17-26.
- [25] Hu, H., Wu, Q., Han, S., Zhang, Z. (2020). Coordination of dual-channel supply chain with perfect product considering sales effort, *Advances in Production Engineering & Management*, Vol. 15, No. 2, 192-203, doi: [10.14743/apem2020.2.358](https://doi.org/10.14743/apem2020.2.358).
- [26] Cheng, H., Su, Y., Yan, J., Wang, X., Li, M. (2019). The incentive model in supply chain with trade credit and default risk, *Complexity*, No. 6, Article ID 5909785, doi: [10.1155/2019/5909785](https://doi.org/10.1155/2019/5909785).
- [27] Deng, S.-M., Wang, T., Wei, D.-J., Ma, S.-H. (2015). Customer satisfaction incentives in decentralized service chains, *Journal of Management Sciences in China*, Vol. 18, No. 8, 12-19.
- [28] Hu, H., Wu, Q., Zhang, Z., Han, S. (2019). Effect of the manufacturer quality inspection policy on the supply chain decision-making and profits, *Advances in Production Engineering & Management*, Vol. 14, No. 4, 472-482, doi: [10.14743/apem2019.4.342](https://doi.org/10.14743/apem2019.4.342).
- [29] Deng, S., Wang, T., Chang, X. (2018). Customer satisfaction incentives with budget constraints, *International Transactions in Operational Research*, Vol. 25, No. 6, 1973-1995, doi: [10.1111/itor.12259](https://doi.org/10.1111/itor.12259).

Testing of novel nano gold ink for inkjet printing

Rudolf, R.^{a,b,*}, Majerič, P.^{a,b}, Golub, D.^a, Tiyyagura, H.R.^a

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

^bZlatarna Celje d.o.o., Celje, Slovenia

ABSTRACT

Gold nanoparticles (GNPs) were synthesised by the Ultrasonic Spray Pyrolysis (USP) process and collected in deionised water with the addition of a stabiliser, i.e. PVP (0.1 wt.%). With the use of a rotary evaporator, a highly concentrated GNPs' suspension was achieved (600 ppm concentration of GNPs), which was used directly as novel nano gold ink for inkjet printing. The physical and chemical characteristics of such prepared nano gold ink were explained in detail by the use of Zeta (ζ) Potential, ATR-FTIR spectroscopy, UV/VIS spectroscopy, and nanoparticle size was identified through SEM. With nano gold ink the chosen pattern was printed onto photo paper, which was characterised for confirming the presence of gold with optical and SEM/EDX observations. The observations revealed that the tested printed nano gold ink on the paper provided a new route for the fabrication of paper-based electrochemical immunosensors, colorimetric sensors and nano-metallic biomedical sensors.

© 2020 CPE, University of Maribor. All rights reserved.

ARTICLE INFO

Keywords:

Inkjet printing;
Nano gold ink;
Gold nanoparticles;
Characterisation;
Paper-based sensor

*Corresponding author:

rebeka.rudolf@um.si
(Rudolf, R.)

Article history:

Received 7 August 2020

Revised 9 September 2020

Accepted 14 September 2020

1. Introduction

Metallic nanoparticles (MNP) have completely different properties like electrical, mechanical, optical and magnetic as their identical bulk materials [1-3]. Based on this, they can be used for various types of applications, such as in the medical and electronic industries [4-6], as well as increasingly in the food industry [7]. The preparation of their inks is envisaged as one of the growing applications of MNP [8]. The metallic nanoparticle ink can be prepared as various types of ink, namely, single element, alloy metallic, metallic oxide and core shell bimetallic nanoparticle inks [9]. These metallic inks, available on the present market, are prepared from Silver, Copper, Gold, Aluminium, Cobalt, Zinc, Palladium, Nickel and Platinum. Nowadays, many conventional methods are being used for the printing of MNP inks on different substrates for different applications [10,11]. Inkjet printing is, namely, a rapidly evolving technology of loading functional inks onto various substrates, and also one of the excellent approaches for the printing of MNP inks, as it deposits the required patterns on the surface with accurate and non-contact writing [12-14]. The described inkjet printing method is operated at room temperature and pressure [8], controlled by data from images or a pattern in conjunction with a computer, and transferring the ink in the form of microdroplets onto the chosen substrate. Inkjet printing has high precision, as well as control of homogeneous microdroplets' loading, and the main advantage is that we can print complex schemes without chemical waste after the printing process, which makes the process environmentally friendly and economical [15].

In the last decade, inks composed of gold nanoparticles (GNPs) – so-called nano gold inks – had a lot of attention towards printed electronics due to the GNPs' properties, like high thermal conductivity, excellent resistance to oxidation and tuneable optical properties [16]. Based on

this, nano gold inks play an important role in the fabrication of printed electronics, biosensing applications (e.g. microelectrode arrays) and electrochemical sensors [17]. The recent literature shows that the preparation of nano gold inks and ink formulations were applied by different approaches [18-21], but these approaches are not beneficial for the production of larger quantities of nano gold ink. The currently available nano gold inks on the market are limited, and cost ineffective when compared to other MNP inks. The aim of this research work was the preparation of novel nano gold ink for the inkjet printing process. The synthesised gold ink presented in this work is cost-effective for the fabrication of paper-based sensor applications.

The rest of this article is arranged as follows: Section 2 describes the synthesis and inkjet printing of GNPs, with the carried out comprehensive characterisation of nano gold ink, including physical and chemical properties, by using different techniques like Zeta (ζ) Potential measurements, ATR-FTIR, UV-VIS spectroscopy, and SEM/EDX observations. Section 3 discusses the results of preliminary testing of printing the patterns by using a Fujifilm Dimatix DMP-2831 inkjet printer. The printed patterns were characterised through Optical Microscopy and SEM/EDX observations. The conclusions and future perspectives are presented in section 4.

2. Materials, methods and design of experiments

2.1 Materials

Synthesis of GNPs

GNPs were synthesised with an Ultrasonic Spray Pyrolysis (USP) device located in Zlatarna Celje d.o.o. (Slovenia), Fig. 1. The USP is composed of an ultrasonic generator (with $f = 2,4$ MHz), a reactor furnace (with 3 temperature zones) and a system for nanoparticle collection (3 collecting bottles). In the USP, with the ultrasound, the precursor solution with the dissolved material is dispersed into droplets. These droplets are then transported with a carrier gas to a high temperature reactor, where the target material inside the droplet is decomposed chemically via pyrolysis, and nanoparticles of pure elements are formed. In this study, the selected raw material for preparing Au-precursor solutions was Au acetate salt (AuAc, gold (III) acetate ($\text{Au}(\text{CH}_3\text{COO})_3$), Alfa Aesar), which was dissolved in deionised water and hydrochloric acid (HCl, 37 %, Sigma Aldrich) with a final concentration of Au 1 g/L [22-26]. The pH value of the prepared solution was about 1-2. The high acidity of this solution may be unfavourable for some of the USP elements during synthesis, as observed from previous practical experience. In order to increase the pH value to 5-6, the solution was stirred magnetically and pellets of sodium hydroxide (NaOH, Fisher Chemicals) were added, obtaining a clear yellow solution, suitable for use with USP. The following parameters were used in the USP synthesis: The flow of carrier gas N_2 was 4 L/min, the flow of reduction gas H_2 was 2 L/min, reactor temperatures were $T_1 = 120$ °C, $T_2 = 400$ °C, $T_3 = 400$ °C, the collection medium was ethanol with stabiliser Polyvinylpyrrolidone (PVP) with a concentration of 2.5 g/L. With USP the formed GNPs were collected in the described collection system.

Preparation of nano gold ink

The nano gold ink was prepared directly from GNPs ethanol/PVP suspension which was, after USP synthesis completion, immediately subjected to a concentration process by using a rotary evaporator. The parameters of the rotary evaporation process were: Rotation: 240 rpm, Pressure: 40 mBar, Bath temperature: 40 °C, Initial volume: 250 mL, Distillation time: 45 min, Final volume: 5 mL. The concentration of GNPs in suspension was measured with Inductively Coupled Plasma-Mass Spectrometry (ICP-MS). The spectrometer used was an HP, Agilent 7500 CE, equipped with a collision cell (Santa Clara, CA, USA). The following conditions for ICP-MS were used: The power was 1.5 kW, Nebuliser-Meinhard, plasma gas flow was 15 L/min, nebuliser gas flow was 0.85, make up gas flow was 0.28 L/min, and reaction gas flow was 4.0 mL/min. The instrument was calibrated with matrix matched calibration solutions. The relative measurement uncertainty was estimated as ± 3 %. The concentration of GNPs in highly concentrated suspension as novel nano gold ink was 600 ppm.



Fig. 1 Ultrasonic Spray Pyrolysis (USP) device located in Zlatarna Celje d.o.o.

Inkjet printing of nano gold ink

The Inkjet printer used for the printing of nano gold ink, was the Fujifilm Dimatix DMP-2831 (FUJIFILM Dimatix, USA) with a piezoelectric Inkjet cartridge. The printing of patterns on a substrate using prepared nano gold ink was under the following parameters: Waveform: low-viscosity (the uniformity of the droplets was the best); Jetting Voltage for all nozzles was 14 V; Tickle control was 23 kHz; Head Angle was 9.5°; Cartridge temperature was 23.5 °C and the Printing plate temperature was 24.5 °C. The print pattern was composed of vertical lines with a thickness of 3 mm and a length of 60 mm. The selected pattern was printed on a commercially obtained glossy photo paper. The pattern was printed in 15 layers in order to ensure a more continuous distribution of GNPs from the nano gold ink.

2.2 Characterisation of nano gold ink and GNPs

Zeta (ζ) Potential measurement

The Zeta (ζ) Potential of nano gold ink, using the Dynamic Light Scattering (DLS) technique, was measured with the Malvern Zetasizer Nano ZS (Malvern Panalytical, UK), and a folded capillary zeta cell. The selected measurement parameters were: Refractive Index (R.I.) for the gold particles was 0.2, absorbance was 3.32, dispersant was ethanol, temperature was 25 °C, R.I. for ethanol was 1.36, viscosity was 1.10 cP, dielectric constant was 22.4.

ATR-FTIR spectroscopy

Attenuated Total Reflectance – Fourier Transform Infrared Spectroscopy (ATR-FTIR) analysis of the nano gold ink and pure PVP – as control – was performed with a Perkin-Elmer FTIR Spectrophotometer and a Golden Gate Attenuated Total Reflection attachment with a diamond crystal. The ATR-FTIR spectra were accumulated within 16 scans at a resolution of 4 cm^{-1} within a range of 4000 cm^{-1} to 650 cm^{-1} .

UV/VIS spectroscopy

The UV/VIS absorption of nano gold ink was measured with a Tecan Infinite M200 UV/VIS Spectrophotometer (Tecan, Austria), using a quartz cuvette. The absorbance measurements were made over the wavelength range of 300-700 nm, with no. flashes = 5x and time per measure = 20 ms.

2.3 Optical and SEM/EDX observations

Optical microscopy

The Inkjet printed layers were examined on the Nikon EPIPHOT 300 light microscope (Nikon, Japan).

SEM microscopy/EDX analysis

A Scanning Electron Microscope (SEM), Sirion 400 NC (FEI, USA) with an Energy-Dispersive X-ray spectroscope (EDX) INCA 350 (Oxford Instruments, UK), was used for the SEM investigations of nano gold ink and printed patterns. Droplets of the nano gold ink were put on SEM holders with conductive carbon adhesive tape and left to dry under vacuum, while the printed patterns were located directly on the SEM holder with conductive carbon adhesive tape without any additional treatment.

EDX was used for the determination of qualitative and semi-quantitative chemical composition.

GNPs' size measurements

GNPs' size measurements in nano gold ink were performed from SEM micrographs. The sizes were measured with the microscope software and with manual measurements from the SEM micrographs, measured with the ImageJ analysis software. The particle size distributions show the manual measurements from the ImageJ software. The GNPs' size distributions were made from 1,000 nanoparticle measurements for each measurement sample. Two types of GNPs' measurements were done: Firstly, on the prepared nano gold ink before filtration, and secondly after manual filtration of nano gold ink through the injection filters with 0.1 μm pore sizes. These measurements were made to determine if the ink was adequate to prevent clogging of GNPs in the nozzle during printing.

3. Results and discussion

3.1 Zeta (ζ) potential measurement

Zeta (ζ) Potential measurement is a significant characterisation technique to determine the surface charge of GNPs in nano gold inks, where the ζ potential can be employed for understanding the physical stability of GNPs [27]. It is generally considered that the GNPs with high negative or positive ζ potential are electrically stabilised, while GNPs with low ζ potentials tend to coagulate [28]. There are many other factors, such as the presence of stabilisers, that affect the physical stability of nano gold inks. The measured ζ potential of GNPs in the studied nano gold ink was -1.89 mV, and the ζ potential distribution can be observed from Fig. 2. A negative ζ potential indicates that GNPs were negatively charged, and this is most likely because GNPs are capped with the PVP stabiliser, which lowers the magnitude of ζ potential. The GNPs capped with PVP stabiliser remained dispersed in the nano gold inks most likely due to the steric hindrance created by the large PVP layer coating on the surface of the GNPs [29].

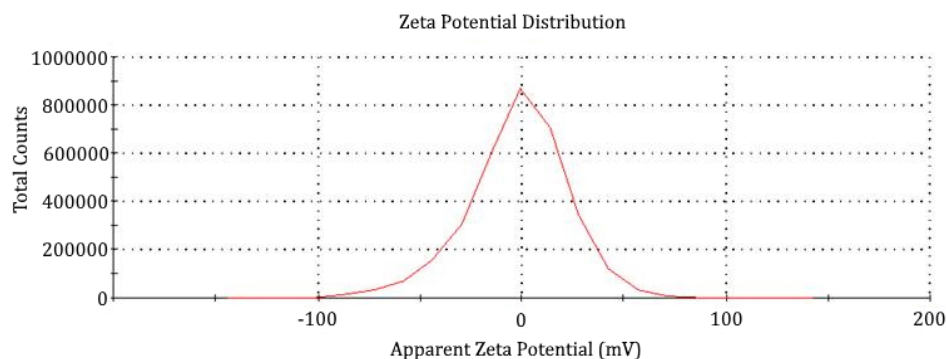


Fig. 2 ζ potential distribution of gold ink

3.2 ATR-FTIR spectroscopy

ATR-FTIR Spectroscopy was used for the determination of PVP stabilised GNPs in nano gold ink – the FTIR spectra of pure PVP and GNPs in nano gold ink are shown in Fig. 3. The FTIR spectra for pure PVP shows characteristic absorption peaks for amide N-H stretch at 3398 cm^{-1} , C-N vibration at 1279 cm^{-1} , C=O stretching at 1667 cm^{-1} and typical peaks for the pyrrolidinyl group of PVP at 1475 cm^{-1} and 1432 cm^{-1} . The characteristic peaks of pure PVP also existed in PVP stabilised nano gold ink, indicating that the adsorption of PVP molecules onto the GNPs' surface was successful [29,30]. The successful functionalisation, most likely via intermolecular hydrogen bonding, was also confirmed with an absorption peak shift of C=O stretching from 1667 cm^{-1} to 1643 cm^{-1} .

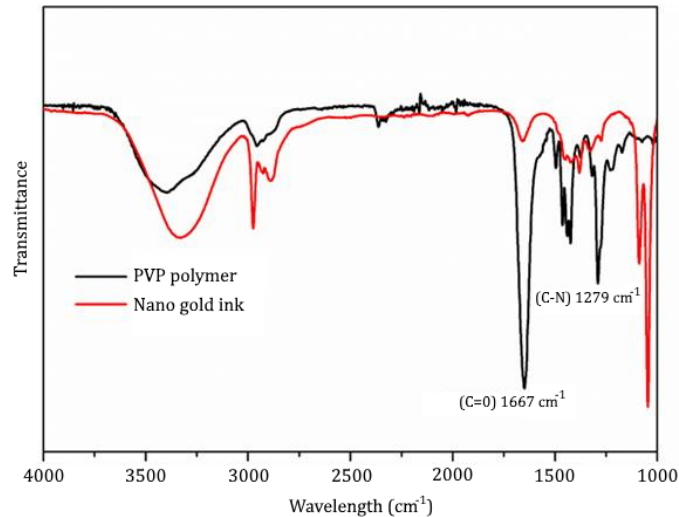


Fig. 3 FTIR spectra of pure PVP polymer (black) and nano gold ink (red)

3.3 UV-VIS spectroscopy

UV-VIS spectroscopy was used for evaluation of the optical and structural properties of nano gold ink, i.e. investigation of the interactions between nano gold ink with different electromagnetic waves. GNPs have unique optical properties, along with a property known as Surface Plasmon Resonance (SPR) [31-32]. With SPR, the absorption of a specific wavelength causes the fluctuation of electrons on the GNPs' surface. SPR is strongly dependent on the GNPs' size, shape and their agglomeration state. It is known that GNPs display a single absorption peak in the visible range between 510-550 nm, and due to the GNPs' size variations, GNPs inks have different colouration [33]. The absorbance spectra of the prepared nano gold ink can be observed from Fig. 4. It was discovered that the maximum absorption band is at $\sim 538\text{ nm}$, which indicates the stable state of the prepared nano gold ink.

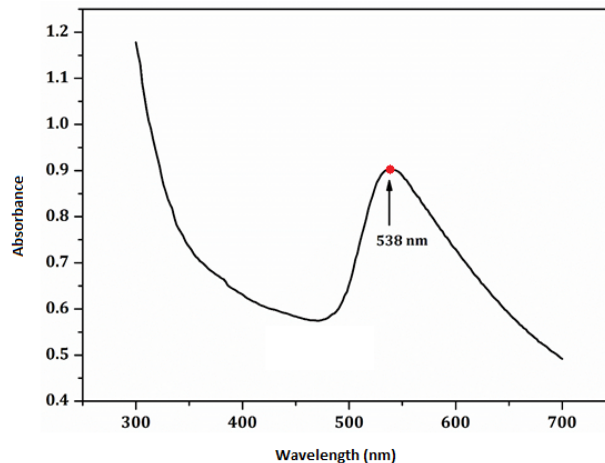


Fig. 4 UV-VIS spectra of nano gold ink

3.4 GNPs' size measurement

GNPs' size measurement was performed using SEM micrographs, firstly on the prepared nano gold ink before filtration (Fig. 5a), and secondly after manual filtration of nano gold ink through injection filters with 0.1 μm pore sizes (Fig. 5b).

The calculated mean GNPs' size in the nano gold ink before filtration was 46.2 nm, with a maximum measured particle size of 332 nm. The smaller GNPs had a mostly spherical shape, while the larger ones had more irregular shapes. After filtration, the mean GNPs size dropped to 16.7 nm. The minimum measured GNPs' size in the filtrated nano gold ink was 5.7 nm, and the maximum particle size was 38 nm. The GNPs' size measurements confirmed that the filtration had removed larger GNPs and clusters of agglomerated GNPs successfully from the prepared nano gold ink, making the ink usable for the printing without clogging of the cartridge nozzles. Mostly spherical and some irregular particles remained in the nano gold ink after filtration, as observed from Fig. 5.

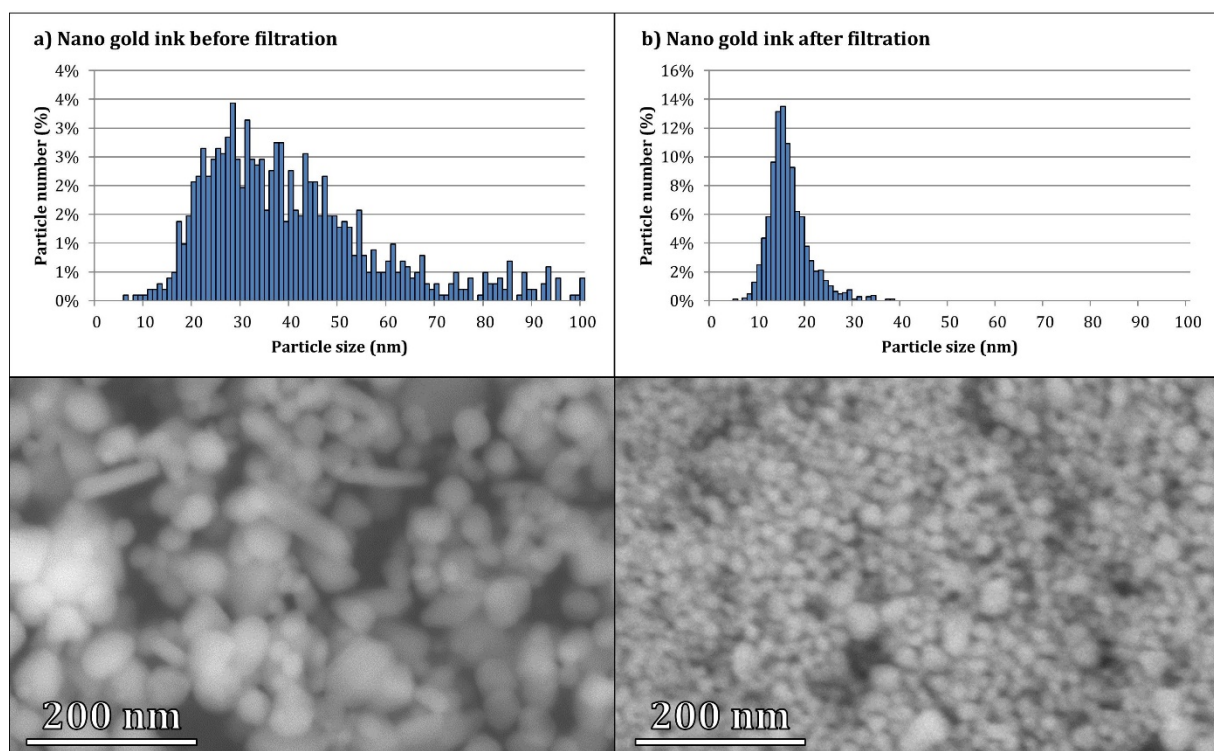


Fig. 5 GNPs' size distributions: a) before and b) after filtration of nano gold ink

3.5 Inkjet printing of nano gold ink

The prepared nano gold ink was injected into the reservoir by means of a syringe and a special metal needle. After the reservoir was filled, a print head was attached to the reservoir, and the assembled cartridge was inserted in the appropriate location in the Inkjet printer. The cartridges used for the printing were Dimatix brand (DMP DMC-11610), with the maximum reservoir volume of 2 mL. The suitable cartridge was first selected in the computer programme. The print head was cleaned by leaking nano gold ink through the nozzles, to make sure that no nozzles were clogged. After the print head was cleaned, a substrate was inserted into the printer and was adhered onto the printing plate at the edges. This prevents the substrate from moving during printing. Before printing, the nano gold ink flow through the nozzles was checked again. With the programme, the individual nozzle and all the nozzles together were checked (Fig. 6a), to ensure that the nano gold ink droplets were homogeneous and uniform. This was achieved with the modification of the ink waveform (Fig. 6b) and with changing the ink droplet parameters.

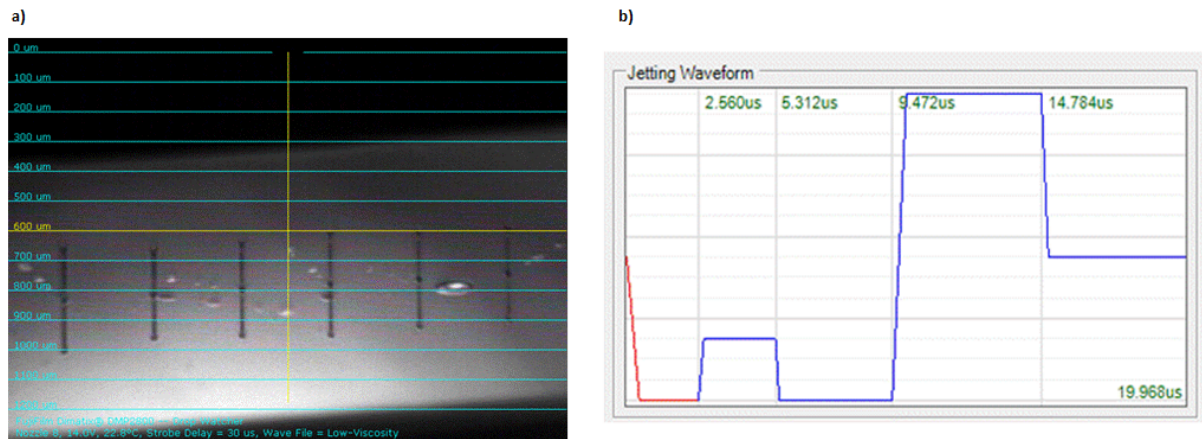


Fig. 6 a) Image of ink jetting through nozzles at the applied voltage of 14 V, which is the threshold for proper drop formation and breakage using the wavefo; b) waveform customised for this study

3.6 Optical microscopy

The nano gold ink Inkjet printing performance of vertical lines printed onto photo paper was observed by optical microscope, as shown in Fig. 7a. It can be seen that the printed line widths between each line do not correspond, most likely due to the overlaying between the spreading of the droplets. The line width becomes smaller with increasing in the drop spacing, which is due to decreasing in the overlapping of the nano gold ink droplets. The measured average printed line width was $\sim 130 \mu\text{m}$.

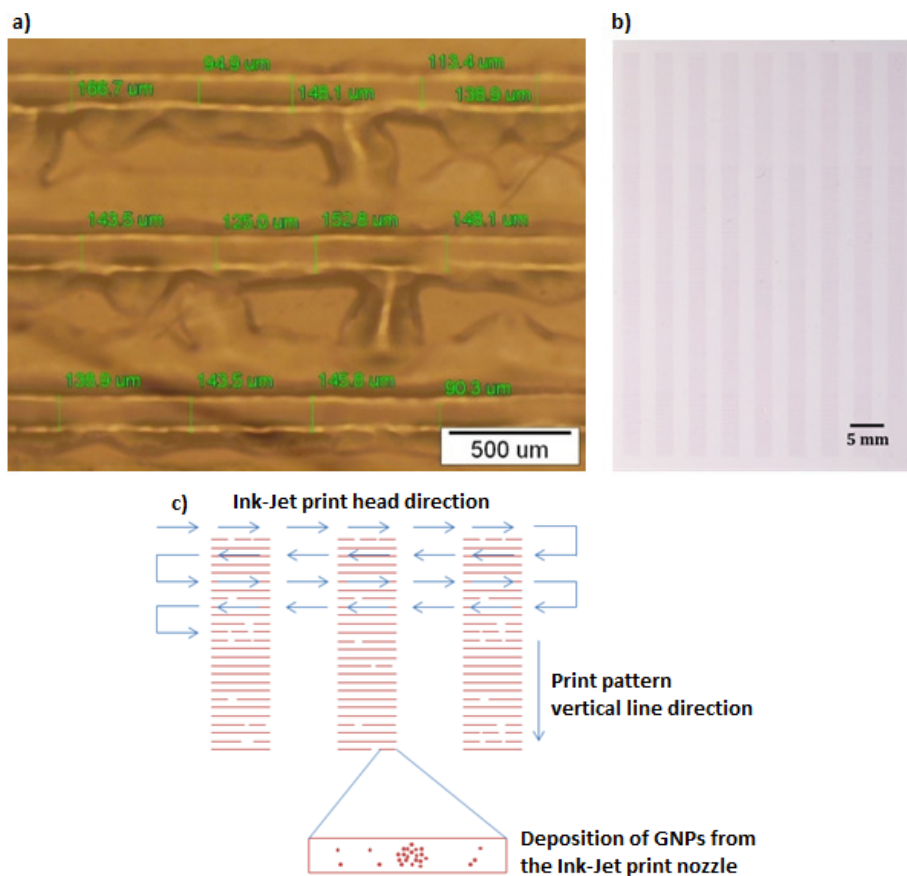


Fig. 7 Optical microscope image of nano gold ink Inkjet printing performance: a) Vertical lines on photo paper; b) Macro image of printed pattern; c) Schematic representation of the pattern printing directions

3.7 SEM microscopy/EDX analysis

A typical SEM micrograph of a nano gold ink printed pattern on pure photo paper is shown in Fig. 8a. The gold printed pattern was investigated thoroughly for determination of the GNPs' distribution (Fig. 8b). The GNPs were clustered together to a greater extent in the optimal region (Fig. 8c). EDX analysis confirmed that the printing with nano gold ink was successful, as the chemical composition of the clusters showed more than 40.1 % of Au (Fig. 8d). The elements C, O, Al and Si were originating from the printing substrate – a commercially obtained glossy photo paper.

The SEM analysed printed pattern does not show a continuous layer of GNPs, but rather visible printing lines, which contain clusters of GNPs and discrete GNPs. These initial results identify that the printing parameters would have to be modified for printing of a more continuous layer of GNPs: Print nozzle size, rheological properties of the nano gold ink (surface tension and viscosity), GNPs' concentration in the nano gold ink, GNPs' sizes, droplet spacing, print overlapping and printing direction [34]. There are some limitations for these parameters that also need to be considered for printing of a continuous layer.

The print nozzle size, along with the rheological properties of the nano gold ink, determines the ink flow for printing. These properties need to keep the nano gold ink inside the nozzle without the ink flowing freely from the nozzle when the printing is stopped and provide the flow of the ink and droplet generation when the Inkjet is activated. Stabilisers for GNPs may alter the rheological properties of the nano gold ink, which is something to consider during nano gold ink formulation.

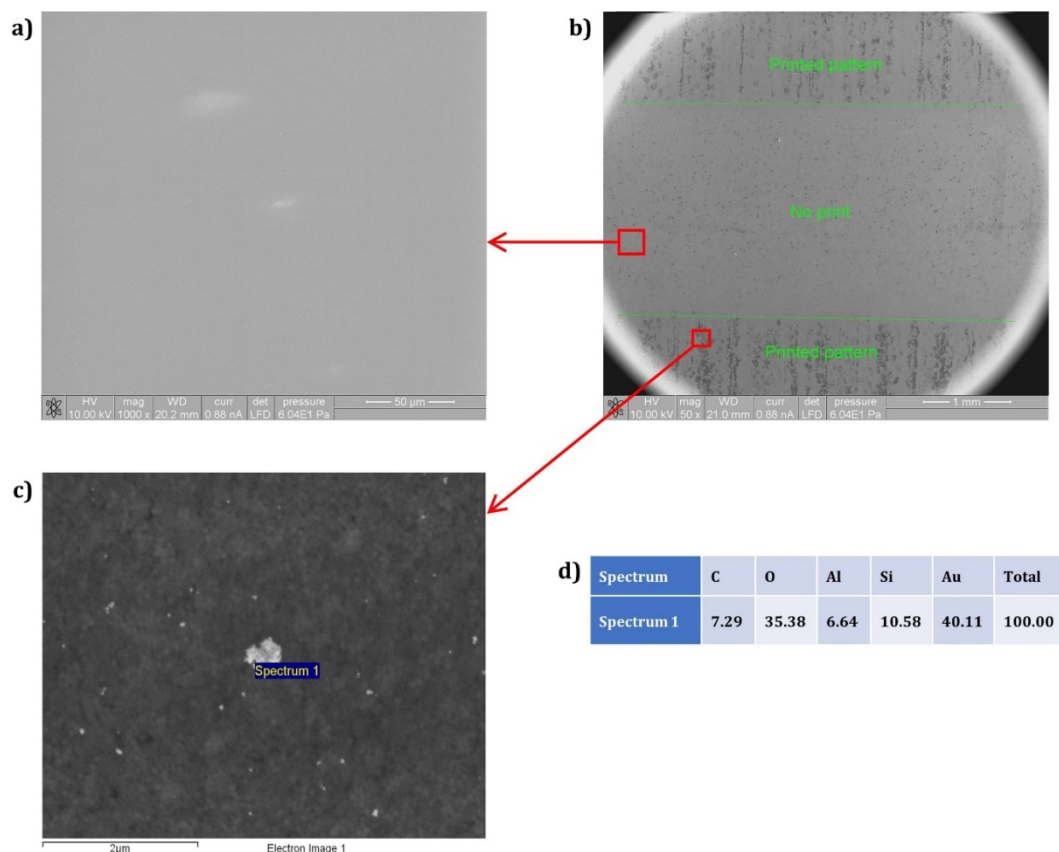


Fig. 8 SEM micrograph of Inkjet printed GNPs with corresponding EDX analysis: a) Pure photo paper; b) Printed pattern with the space between the lines; c) GNPs' distribution; d) EDX analysis

The available nozzle sizes also limit the GNPs' sizes in nano gold ink, which can be used for successful printing. Our experimental results show that GNPs of size around 10 nm are still capable of printing, while larger sizes clog the printing nozzle, making nano gold ink unusable. Since the GNPs in nano gold ink, obtained with the USP technique had a broad size distribution, filtering of nano gold ink was mandatory for printing with an Inkjet printer. Another point to consider is the concentration of GNPs in the nano gold ink and their stability. Using very high concentrations of GNPs (> 600 ppm) leads to unstable dispersions (depending on the stabiliser used), resulting in agglomeration, rendering the nano gold ink unusable. Very high concentrations may also lead to unwanted Inkjet behaviour, interfering with ink droplet sizes and droplet generation intervals. The distribution of GNPs inside the droplet is also a factor, affecting the printed pattern. These combined nano gold ink properties may result in various forms of the printed patterns. One such example is the "coffee-ring effect" [9], where most of the GNPs would be clustered on the droplet outer diameter, while the inside of the printed droplet has a disproportionately low number of GNPs.

Possibilities for improving the printed pattern may be by modifying the droplet spacing, print overlapping and printing direction, in order to produce more favourable results for printing a more continuous layer of GNPs. Several experiments need to be performed in order to find the optimal balance of nano gold ink properties, GNPs' properties and printing parameters. Here, it was shown that USP is capable of providing GNPs to be used in Inkjet printing, while the discussed printing specifications depend on the application of the printed patterns.

Using a printed layer of discretely dispersed GNPs on glass or other transparent substrates produces a spectrally selective surface which blocks infrared radiation due to the plasmon resonance effect. This makes these printed GNPs usable in specialised optical applications, such as light filters or logic gates [35]. An example of exploiting this property is also usage in energy-efficient windows. Note, however, that, nowadays, some other materials are better suited for this particular task [35]. The catalytic properties of GNPs may also be used for CO oxidation into CO₂, using these GNPs in gas masks and some chemical processes, as well as for the removal of CO traces [35]. They may also be used as decorative inks, or as a colouring additive in niche glassware and ceramics. The successful printing of GNPs produced by USP thus shows the possibility of utilising these GNPs for the fabrication of specialised products.

4. Conclusion

Nano gold ink prepared directly from a highly concentrated GNPs' suspension synthesised through USP synthesis, was investigated in the present work. Results of different characterisation showed that the high stability of the nano gold ink was confirmed. ζ potential measurements showed that GNPs were negatively charged, due to the capping of GNPs with PVP stabiliser, and the successful functionalisation of PVP was confirmed with ATR-FTIR. The average size of GNPs in nano gold ink was 16.7 nms after manual filtration through injection filters with 0.1 μ m pore sizes. From this case study, we can conclude that the nano gold ink printed patterns are highly stable, while the printing specifics can be modified for different applications. The preparation of nano gold ink with USP and the selected synthesis parameters required filtration of the nanoparticles in order to prevent clogging of the inkjet printing head. The given initial results also show a non-continuous layer of printed GNPs, indicating that the ink and printing parameters would have to be modified for more favourable results. The presented results demonstrate that the preparation of nano gold inks with the USP technique can contribute in the fabrication of specialised products, useful for different types of technological applications such as electrochemical and nano-metallic biosensors.

Considering the successful synthesis of gold nano ink, the preparation of lower concentrations of GNPs with sizes below 10 nm, and an evaluation of their properties, printing characteristics and real-life analysis for their application in paper-based sensors, will be tested in future studies.

Acknowledgement

This research was funded by the Ministry of Education, Science and Sport, Republic of Slovenia and the European Union, The European Regional Development Fund (ERDF), Early research careers 2.1. We would like to thank Lidija Rozman Zorko for performing the SEM analysis on the GNPs.

References

- [1] Parab, H.J., Chen, H.M., Bagkar, N.C., Liu, R.-S., Hwu, Y.-K., Tsai, D.P. (2010). Approaches to the synthesis and characterization of spherical and anisotropic noble metal nanomaterials, *Nanotechnologies for the Life Sciences*, Vol. 1, 405-459, doi: [10.1002/9783527610419.ntls0133](https://doi.org/10.1002/9783527610419.ntls0133).
- [2] Singh, A., Gautam, P.K., Verma, A., Singh, V., Shivapriya, P.M., Shivalkar, S., Sahoo, A.K., Samanta, S.K. (2020). Green synthesis of metallic nanoparticles as effective alternatives to treat antibiotics resistant bacterial infections: A review, *Biotechnology Reports*, Vol. 25, Article No. e00427, doi: [10.1016/j.btre.2020.e00427](https://doi.org/10.1016/j.btre.2020.e00427).
- [3] Barberio, M., Giusepponi, S., Vallières, S., Scisciò, M., Celino, M., Antici, P. (2020). Ultra-fast high-precision metallic nanoparticle synthesis using laser-accelerated protons, *Scientific Reports*, Vol. 10, No. 1, Article No. 9570, doi: [10.1038/s41598-020-65282-9](https://doi.org/10.1038/s41598-020-65282-9).
- [4] Yaqoob, S.B., Adnan, R., Rameez Khan, R.M., Rashid, M. (2020). Gold, silver, and palladium nanoparticles: A chemical tool for biomedical applications, *Frontiers in Chemistry*, Vol. 8, Article No. 376, doi: [10.3389/fchem.2020.00376](https://doi.org/10.3389/fchem.2020.00376).
- [5] Rudolf, R., Popović, D., Tomić, S., Bobovnik, R., Lazić, V., Majerič, P., Anžel, I., Čolić, M. (2020). Microstructure characterisation and identification of the mechanical and functional properties of a new PMMA-ZnO composite, *Materials*, Vol. 13, No. 12, Article No. 2717, doi: [10.3390/ma13122717](https://doi.org/10.3390/ma13122717).
- [6] Kuppe, C., Rusimova, K.R., Ohnoutek, L., Slavov, D., Valev, V.K. (2020). "Hot" in plasmonics: Temperature-related concepts and applications of metal nanostructures, *Advanced Optical Materials*, Vol. 8, No. 1, Article No. 1901166, doi: [10.1002/adom.201901166](https://doi.org/10.1002/adom.201901166).
- [7] Shafiq, M., Anjum, S., Hano, C., Anjum, I., Abbasi, B.H. (2020). An overview of the applications of nanomaterials and nanodevices in the food industry, *Foods*, Vol. 9, No. 2, Article No.148, doi: [10.3390/foods9020148](https://doi.org/10.3390/foods9020148).
- [8] Maleki, H., Bertola, V. (2020). Recent advances and prospects of inkjet printing in heterogeneous catalysis, *Catalysis Science & Technology*, Vol. 10, No. 10, 3140-3159, doi: [10.1039/D0CY00040J](https://doi.org/10.1039/D0CY00040J).
- [9] Tan, H.W., An, J., Chua, C.K., Tran, T. (2019). Metallic nanoparticle inks for 3D printing of electronics, *Advanced Electronic Materials*, Vol. 5, No. 5, Article No. 1800831, doi: [10.1002/aelm.201800831](https://doi.org/10.1002/aelm.201800831).
- [10] Cano-Raya, C., Denchev, Z.Z., Cruz, S.F., Viana, J.C. (2019). Chemistry of solid metal-based inks and pastes for printed electronics – A review, *Applied Materials Today*, Vol. 15, 416-430, doi: [10.1016/j.apmt.2019.02.012](https://doi.org/10.1016/j.apmt.2019.02.012).
- [11] Calvert, P. (2001). Inkjet printing for materials and devices, *Chemistry of Materials*, Vol. 13, No. 10, 3299-3305, doi: [10.1021/cm0101632](https://doi.org/10.1021/cm0101632).
- [12] Derby, B. (2010). Inkjet printing of functional and structural materials: Fluid property requirements, feature stability, and resolution, *Annual Review of Materials Research*, Vol. 40, No. 1, 395-414, doi: [10.1146/annurev-matsci-070909-104502](https://doi.org/10.1146/annurev-matsci-070909-104502).
- [13] Cruz, B., Albrecht, A., Eschlwach, P., Biebl, E. (2019). Inkjet printing of metal nanoparticles for green UHF RFID tags, *Advances in Radio Science*, Vol. 17, 119-127, doi: [10.5194/ars-17-119-2019](https://doi.org/10.5194/ars-17-119-2019).
- [14] Karim, N., Afroj, S., Tan, S., Novoselov, K.S., Yeates, S.G. (2019). All inkjet-printed graphene-silver composite ink on textiles for highly conductive wearable electronics applications, *Scientific Reports*, Vol. 9, No. 1, Article No. 8035, doi: [10.1038/s41598-019-44420-y](https://doi.org/10.1038/s41598-019-44420-y).
- [15] Soleimani-Gorgani, A. (2016). 14 – Inkjet Printing, In Izdebska, J., Thomas, S. (eds.), *Printing on Polymers*, William Andrew Publishing, New York, USA, 231-246, doi: [10.1016/B978-0-323-37468-2.00014-2](https://doi.org/10.1016/B978-0-323-37468-2.00014-2).
- [16] Chu, J.-H., Joo, S.-J., Kim, H.-S. (2019). Development of a via-hole connection process via intense pulsed light sintering with Cu micro/Ag nano-hybrid ink for a multi-layered flexible printed circuit board, *Thin Solid Films*, Vol. 680, 1-11, doi: [10.1016/j.tsf.2019.04.015](https://doi.org/10.1016/j.tsf.2019.04.015).
- [17] Bishop, P.T., Ashfield, L.J., Berzins, A., Boardman, A., Buche, V., Cookson, J., Gordon, R.J., Salcianu, C., Sutton, P.A. (2010). Printed gold for electronic applications, *Gold Bulletin*, Vol. 43, No. 3, 181-188, doi: [10.1007/BF03214985](https://doi.org/10.1007/BF03214985).
- [18] Deng, M., Zhang, X., Zhang, Z., Xin, Z., Song, Y. (2014). A gold nanoparticle ink suitable for the fabrication of electrochemical electrode by inkjet printing, *Journal of Nanoscience and Nanotechnology*, Vol. 14, No. 7, 5114-5119, doi: [10.1166/jnn.2014.7208](https://doi.org/10.1166/jnn.2014.7208).
- [19] Su, C.-H., Chiu, H.-L., Chen, Y.-C., Yesilmen, M., Schulz, F., Ketelsen, B., Vossmeier, T., Liao, Y.-C. (2019). Highly responsive PEG/gold nanoparticle thin-film humidity sensor via inkjet printing technology, *Langmuir*, Vol. 35, No. 9, 3256-3264, doi: [10.1021/acs.langmuir.8b03433](https://doi.org/10.1021/acs.langmuir.8b03433).
- [20] Bacalzo Jr., N.P., Go, L.P., Querebillo, C.J., Hildebrandt, P., Limpoco, F.T., Enriquez, E.P. (2018). Controlled microwave-hydrolyzed starch as a stabilizer for green formulation of aqueous gold nanoparticle ink for flexible printed electronics, *ACS Applied Nano Materials*, Vol. 1, No. 3, 1247-1256, doi: [10.1021/acsnm.7b00379](https://doi.org/10.1021/acsnm.7b00379).
- [21] Pavithra, M., Muruganand, S., Parthiban, C. (2018). Development of novel paper based electrochemical immunosensor with self-made gold nanoparticle ink and quinone derivate for highly sensitive carcinoembryonic antigen, *Sensors and Actuators B: Chemical*, Vol. 257, 496-503, doi: [10.1016/j.snb.2017.10.177](https://doi.org/10.1016/j.snb.2017.10.177).

- [22] Golub, D., Ivanič, A., Majerič, P., Tiyyagura, H.R., Anžel, I., Rudolf, R. (2019). Synthesis of colloidal Au nanoparticles through ultrasonic spray pyrolysis and their use in the preparation of polyacrylate-AuNPs' composites, *Materials*, Vol. 12, No. 22, Article No. 3775, doi: [10.3390/ma12223775](https://doi.org/10.3390/ma12223775).
- [23] Rudolf, R., Shariq, M., Veselinovic, V., Adamovic, T., Bobovnik, R., Kargl, R., Majeric, P. (2018). Synthesis of gold nanoparticles through ultrasonic spray pyrolysis and its application in printed electronics, *Contemporary Materials*, Vol. 9, No. 1, 106-112.
- [24] Shariq, M., Marić, N., Gorše, G.K., Kargl, R., Rudolf, R. (2018). Synthesis of gold nanoparticles with ultrasonic spray pyrolysis and its feasibility for ink-jet printing on paper, *Micro and Nanosystems*, Vol. 10, No. 2, 102-109, doi: [10.2174/1876402910666180802113859](https://doi.org/10.2174/1876402910666180802113859).
- [25] Shariq, M., Friedrich, B., Budic, B., Hodnik, N., Ruiz-Zepeda, F., Majerič, P., Rudolf, R. (2018). Successful synthesis of gold nanoparticles through ultrasonic spray pyrolysis from a gold(III) nitrate precursor and their interaction with a high electron beam, *ChemistryOpen*, Vol. 7, No. 7, 533-542, doi: [10.1002/open.201800101](https://doi.org/10.1002/open.201800101).
- [26] Shariq, M., Majerič, P., Friedrich, B., Budic, B., Jenko, D., Dixit, A.R., Rudolf, R. (2017). Application of gold(III) acetate as a new precursor for the synthesis of gold nanoparticles in PEG through ultrasonic spray pyrolysis, *Journal of Cluster Science*, Vol. 28, No. 3, 1647-1665, doi: [10.1007/s10876-017-1178-0](https://doi.org/10.1007/s10876-017-1178-0).
- [27] Kumar, A., Dixit, C.K. (2017). 3 - Methods for characterization of nanoparticles, In: Nimesh, S., Chandra, R., Gupta, N., (eds.), *Advances in nanomedicine for the delivery of therapeutic nucleic acids*, Woodhead Publishing, Cambridge, United Kingdom, 43-58, doi: [10.1016/B978-0-08-100557-6.00003-1](https://doi.org/10.1016/B978-0-08-100557-6.00003-1).
- [28] Chirayil, C.J., Abraham, J., Mishra, R.K., George, S.C., Thomas, S. (2017). Chapter 1 – Instrumental techniques for the characterization of nanoparticles, In: Thomas, S., Thomas, R., Zachariah, A.K., Mishra, R.K. (eds.), In: *Thermal and rheological measurement techniques for nanomaterials characterization, micro and nano technologies*, Elsevier, Amsterdam, Netherlands, 1-36, doi: [10.1016/B978-0-323-46139-9.00001-3](https://doi.org/10.1016/B978-0-323-46139-9.00001-3).
- [29] Das, T., Kolli, V., Karmakar, S., Sarkar, N. (2017). Functionalisation of polyvinylpyrrolidone on gold nanoparticles enhances its anti-amyloidogenic propensity towards hen egg white lysozyme, *Biomedicines*, Vol. 5, No. 2, Article No. 19, doi: [10.3390/biomedicines5020019](https://doi.org/10.3390/biomedicines5020019).
- [30] Dhumale, V.A., Gangwar, R.K., Datar, S.S., Sharma, R.B. (2012). Reversible aggregation control of polyvinylpyrrolidone capped gold nanoparticles as a function of pH, *Materials Express*, Vol. 2, No. 4, 311-318, doi: [10.1166/mex.2012.1082](https://doi.org/10.1166/mex.2012.1082).
- [31] Amendola, V., Meneghetti, M. (2009). Size evaluation of gold nanoparticles by UV-vis spectroscopy, *The Journal of Physical Chemistry C*, Vol. 113, No. 11, 4277-4285, doi: [10.1021/jp8082425](https://doi.org/10.1021/jp8082425).
- [32] Tiyyagura, H.R., Majerič, P., Anžel, I., Rudolf, R. (2020). Low-cost synthesis of AuNPs through ultrasonic spray pyrolysis, *Materials Research Express*, Vol. 7, No. 5, Article No. 055017, doi: [10.1088/2053-1591/ab80ea](https://doi.org/10.1088/2053-1591/ab80ea).
- [33] Lin, P.-C., Lin, S., Wang, P.C., Sridhar, R. (2014). Techniques for physicochemical characterization of nanomaterials, *Biotechnology Advances*, Vol. 32, No. 4, 711-726, doi: [10.1016/j.biotechadv.2013.11.006](https://doi.org/10.1016/j.biotechadv.2013.11.006).
- [34] Shariq, M., Chattopadhyaya, S., Rudolf, R., Rai Dixit, A. (2020). Characterization of AuNPs based ink for inkjet printing of low cost paper based sensors, *Materials Letters*, Vol. 264, Article No. 127332, doi: [10.1016/j.matlet.2020.127332](https://doi.org/10.1016/j.matlet.2020.127332).
- [35] Cortie, M. (2011). Technological applications of gold nanoparticles, In: Louis, C., Pluchery, O. (eds.), *Gold nanoparticles for physics, chemistry and biology*, Imperial College Press, London, United Kingdom, 355-377, doi: [10.1142/9781848168077_0013](https://doi.org/10.1142/9781848168077_0013).

Calendar of events

- 21st International Conference and Exhibition on Materials Science, Nanotechnology and Engineering, September 21-22, 2020, Milan, Italy.
- 31st DAAAM International Symposium - Virtual Online Edition, October 21-24, 2020, hosted from Mostar, Bosnia and Herzegovina.
- 4th annual European Simulation and Modelling Conference (ESM 2020), October 21-23, 2020, Toulouse, France.
- 15th International Conference on Green Supply Chain Management Applications, January 18-19, 2021, Rome, Italy.
- International Conference on Manufacturing Models and Cloud Manufacturing ICMMCM, July 15-16, 2021, Stockholm, Sweden.
- 15th International Conference on Industrial Production Methods and Flow Production, August 5-6, 2021, Montreal, Canada.

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. Authors of submitted papers automatically accept a copyright transfer to *Chair of Production Engineering, University of Maribor*. For most up-to-date information on publishing procedure please see the *APEM journal* homepage apem-journal.org.

Submission of papers

A submission must include the corresponding author's complete name, affiliation, address, phone and fax numbers, and e-mail address. All papers for consideration by *Advances in Production Engineering & Management* should be submitted by e-mail to the journal Editor-in-Chief:

Miran Brezocnik, Editor-in-Chief
UNIVERSITY OF MARIBOR
Faculty of Mechanical Engineering
Chair of Production Engineering
Smetanova ulica 17, SI – 2000 Maribor
Slovenia, European Union
E-mail: editor@apem-journal.org

Manuscript preparation

Manuscript should be prepared in *Microsoft Word 2010* (or higher version) word processor. *Word .docx* format is required. Papers on A4 format, single-spaced, typed in one column, using body text font size of 11 pt, should not exceed 12 pages, including abstract, keywords, body text, figures, tables, acknowledgements (if any), references, and appendices (if any). The title of the paper, authors' names, affiliations and headings of the body text should be in *Calibri* font. Body text, figures and tables captions have to be written in *Cambria* font. Mathematical equations and expressions must be set in *Microsoft Word Equation Editor* and written in *Cambria Math* font. For detail instructions on manuscript preparation please see instruction for authors in the *APEM journal* homepage apem-journal.org.

The review process

Every manuscript submitted for possible publication in the *APEM journal* is first briefly reviewed by the editor for general suitability for the journal. Notification of successful submission is sent. After initial screening, and checking by a special plagiarism detection tool, the manuscript is passed on to at least two referees. A double-blind peer review process ensures the content's validity and relevance. Optionally, authors are invited to suggest up to three well-respected experts in the field discussed in the article who might act as reviewers. The review process can take up to eight weeks on average. Based on the comments of the referees, the editor will take a decision about the paper. The following decisions can be made: accepting the paper, reconsidering the paper after changes, or rejecting the paper. Accepted papers may not be offered elsewhere for publication. The editor may, in some circumstances, vary this process at his discretion.

Proofs

Proofs will be sent to the corresponding author and should be returned within 3 days of receipt. Corrections should be restricted to typesetting errors and minor changes.

Offprints

An e-offprint, i.e., a PDF version of the published article, will be sent by e-mail to the corresponding author. Additionally, one complete copy of the journal will be sent free of charge to the corresponding author of the published article.

APEM

journal

Advances in Production Engineering & Management

Chair of Production Engineering (CPE)
University of Maribor
APEM homepage: apem-journal.org

Volume 15 | Number 3 | September 2020 | pp 251-372

Contents

Scope and topics	254
An exact approach for the consistent vehicle routing problem (ConVRP) Barros, L.; Linfati, R.; Escobar, J.W.	255
Decentralized optimization of the flexible production lines Malega, P.; Rudy, V.; Kanász, R.; Gazda, V.	267
Interactive impacts of overconfidence and fairness concern on supply chain performance Zhang, Z.J.; Wang, P.; Wan, M.Y.; Guo, J.H.; Luo, C.L.	277
Computational analysis of cavitation at the tongue of the volute of a centrifugal pump at overload conditions Hu, Q.; Yang, Y.; Cao, W.	295
Optimization of process performance by multiple pentagon fuzzy responses: Case studies of wire-electrical discharge machining and sputtering process Al-Refaie, A.; Lepkova, N.; Abbasi, G.; Bani Domi, G.	307
An integrated CNC system for chatter suppression in turning Jasiewicz, M.; Miqdlicki, K.	318
Simulation-based time evaluation of basic manual assembly tasks Turk, M.; Pipan, M.; Šimic, M.; Herakovič, N.	331
Manufacturer's customer satisfaction incentive plan for duopoly retailers with Cournot or collusion games Hu, H.; Zhang, Z.; Wu, Q.; Han, S.	345
Testing of novel nano gold ink for inkjet printing Rudolf, R.; Majerič, P.; Golub, D.; Tiyyagura, H.R.	358
Calendar of events	369
Notes for contributors	371

Copyright © 2020 CPE. All rights reserved.



apem-journal.org