

ISSN 1854-6250

APEM
journal

Advances in Production Engineering & Management

Volume 18 | Number 4 | December 2023



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



Subsidizer: The journal is subsidized by Slovenian Research Agency



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

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

Published by CPE, University of Maribor.

Advances in Production Engineering & Management is indexed and abstracted in the **WEB OF SCIENCE** (maintained by **Clarivate 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 18 | Number 4 | December 2023 | pp 399–514

Contents

Scope and topics	402
A combined genetic algorithm and A* search algorithm for the electric vehicle routing problem with time windows	403
Wang, D.L.; Ding, A.; Chen, G.L.; Zhang, L.	
Optimizing rock breaking performance: The influence of chamfer on polycrystalline diamond compact (PDC) cutters	417
Ju, P.	
Dynamic price competition market for retailers in the context of consumer learning behavior and supplier competition: Machine learning-enhanced agent-based modeling and simulation	434
Deng, G.F.	
IoT-based Deep Learning Neural Network (DLNN) algorithm for voltage stability control and monitoring of solar power generation	447
Shweta, R.; Sivagnanam, S.; Kumar, K.A.	
Comparing Fault Tree Analysis methods combined with Generalized Grey Relation Analysis: A new approach and case study in the automotive industry	462
Shi, J.L.; Lu, Z.C.; Xu, H.H.; Ren, M.M.; Shu, F.L.	
Optimizing smart manufacturing systems using digital twin	475
Ojstersek, R.; Javernik, A.; Buchmeister, B.	
Incentive modeling analysis in engineering applications and projects with stochastic duration time	486
Zhao, J.; Su, J.F.	
Reduction of surface defects by optimization of casting speed using genetic programming: An industrial case study	501
Kovacic, M.; Zuperl, U.; Gusel, L.; Brezocnik, M.	
Calendar of events	512
Notes for contributors	513

Journal homepage: apem-journal.org

ISSN 1854-6250 (print)

ISSN 1855-6531 (on-line)

Published by CPE, University of Maribor.

Scope and topics

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

Fields of interest include, but are not limited to:

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

A combined genetic algorithm and A* search algorithm for the electric vehicle routing problem with time windows

Wang, D.L.^a, Ding, A.^a, Chen, G.L.^b, Zhang, L.^{a,c,*}

^aSchool of Economics and Management, China University of Geosciences (Beijing), Haidian District, Beijing, P.R. China

^bSchool of Geophysics and Information Technology, China University of Geosciences (Beijing), Haidian District, Beijing, P.R. China

^cFaculty of Business, City University of Macau, Macau, P.R. China

ABSTRACT

With growing environmental concerns, the focus on greenhouse gases (GHG) emissions in transportation has increased, and the combination of smart microgrids and electric vehicles (EVs) brings a new opportunity to solve this problem. Electric vehicle routing problem with time windows (EVRPTW) is an extension of the vehicle routing problem (VRP) problem, which can reach the combination of smart microgrids and EVs precisely by scheduling the EVs. However, the current genetic algorithm (GA) for solving this problem can easily fall into the dilemma of local optimization and slow iteration speed. In this paper, we present an integer hybrid planning model that introduces time of use and area price to enhance realism. We propose the GA-A* algorithm, which combines the A* algorithm and GA to improve global search capability and iteration speed. We conducted experiments on 16 benchmark cases, comparing the GA-A* algorithm with traditional GA and other search algorithms, results demonstrate significant enhancements in searchability and optimal solutions. In addition, we measured the grid load, and the model implements the vehicle-to-grid (V2G) mode, which serves as peak shaving and valley filling by integrating EVs into the grid for energy delivery and exchange through battery swapping. This research, ranging from model optimization to algorithm improvement, is an important step towards solving the EVRPTW problem and improving the environment.

ARTICLE INFO

Keywords:

Vehicle routing problem (VRP);
Electric vehicle;
Optimization;
Time windows;
Spatiotemporal electricity price;
Smart microgrids;
Genetic algorithm (GA);
A* search algorithm;
GA-A* algorithm

*Corresponding author:

zhanglongdragon@hotmail.com
(Zhang, L.)

Article history:

Received 25 September 2023

Revised 17 November 2023

Accepted 21 November 2023



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

1. Introduction

With global warming and environmental problems intensifying, more and more attention has been paid to reducing greenhouse gas emissions in transportation. New energy vehicles have emerged globally as an environmentally friendly and energy-efficient alternative [1], but face challenges in practical applications due to limitations in battery technology and charging facilities [2]. To overcome these challenges, smart microgrids have emerged as a new type of energy management system, which can improve power sharing and power quality improvement and promote the sustainable development of new energy vehicles [3, 4]. Meanwhile, AI plays an important role in solving energy management problems in smart microgrids by applying various optimization methods and developing optimal energy management strategies, which can help to increase energy efficiency, reduce the total cost, and improve power quality [5].

EVRPTW as an optimization problem, aims to efficiently plan the routes of electric vehicles to meet the time window constraints and minimize the total cost or maximize the benefits. It optimizes the routes and charging strategies of electric vehicles to maximize the use of renewable energy and grid power for efficient energy management. By rationally arranging the use of charging stations and charging periods, it can reduce the dependence on the traditional grid and improve energy utilization. Meanwhile, by rationally planning the routes of electric vehicles, EVRPTW can reduce pollutant emissions and carbon footprint. Introducing EVs into smart microgrids and the whole grid, can promote the use of green energy and reduce dependence on oil and other limited resources.

Keskin and Çatay first started working on the EVRPTW problem in 2016 and specifically studied the allowed partial charging problem (EVRPTW-PR), where they formulated the problem as 0-1 mixed integer linear programming, developed an adaptive large neighborhood search (ALNS) algorithm, and used benchmark instances to test those solutions that can effectively find high quality solutions, As the result shows, the partial tolling scheme can substantially improve the routing decisions [6]. Keskin and Çatay reformulated the problem as mixed integer linear programming in 2018, where they built on the original by combining an ALNS method with an exact method equipped with various destroy-repair algorithms to efficiently explore neighborhoods and using CPLEX to strengthen the obtained routes [7]. In 2019, Wang *et al.* constructed a mathematical model to minimize the total cost based on EVRPTW considering time windows and battery swapping stations (2E-EVRPTW-BSS) for the two-stage vehicle path problem of electric vehicles and verified the validity of the model [8]. Gocken and Yaktubay solved the VRPTW problem by applying first clustering and then using genetic algorithm solution for planning. Meanwhile different clustering algorithms are compared and finally the superiority of K-means algorithm for initialising the population is concluded [9]. In 2020, Raeesi and Zografos developed a program utilizing a two-phase hybrid of dynamic programming and integer programming algorithms. The resulting program serves as the cornerstone of a robust, large neighborhood search algorithm, designed for the rapid resolution of instances related to EVRPTW-SMBS [10]. In 2021, Zhu *et al.* applied the elitist genetic algorithm to the EV path problem with a time window [11]. Deng *et al.* studied the EV path problem with a time window and nonlinear charging constraints (EVRPTW-NL) and proposed an improved hybrid algorithm combining an improved differential evolution (IDE) and several heuristics [12]. Bac and Erdem developed a series of neighborhood operators for the EVRPTW problem in the Variable Neighborhood Search (VNS) and Variable Neighborhood Drop (VND) heuristics for the local search process [13]. Lin *et al.* came out with EVRPTW (EVRPTW-TP) under time-varying tariffs, formulating it as an optimal problem, proposing a Lagrangian relaxation method and a mixed-variable neighborhood search/tabu search heuristic to obtain high-quality lower bounds and feasible solutions, respectively [14]. Lin *et al.* proposed an end-to-end deep reinforcement learning framework for solving the EVRPTW [15]. In 2022, Erdelić and Carić implemented the Adaptive Large Neighborhood Search (ALNS) meta-heuristic algorithm, utilizing the ruin-recreate strategy. This algorithm integrates a novel initial solution heuristic, partial search, path removal, and an exact procedure, resulting in the achievement of an optimal layout for charging stations. The results show that ALNS can find 38 new optimal solutions on the benchmark EVRPTW instance and that the advantages and disadvantages of using a partial charging strategy compared to a full charging strategy are evident [16]. In 2022, Niu *et al.* proposed the idea of consumers being able to choose multiple delivery addresses and used a large-scale neighbourhood search algorithm to facilitate further matching of logistics and distribution companies with customer needs [17]. In 2022, Liu *et al.* developed a hybrid Genetic Algorithm that combines the 2-opt algorithm with GA [18]. Ding designed an adaptive particle swarm optimization algorithm for the driving cycle based time window electric vehicle routing problem (EVRPTW-DC) to solve the problem [19]. In 2023, Kumar *et al.* proposed a firefly and ant colony algorithm with a new pad heuristic avoiding local optimums [20]. Kempton and Tomić first proposed electric vehicle-to-grid (V2G) technology, which utilizes large amounts of electric vehicle energy storage through interaction between electric vehicles and the grid to act as a buffer between the grid and renewable energy sources [21].

In 2018, Shao *et al.* solved the EVRP model based on a hybrid genetic algorithm and used the dynamic Dijkstra algorithm to make some improvements to the classical Dijkstra algorithm [22]. In the same year, Wang *et al.* proposed to solve complex multi-objective problems based on a heuristic algorithm (ST-VNSGA) consisting of a variable neighbourhood search method and GA considering the spatiotemporal distance [23]. Zhu *et al.* investigated the path algorithm, and then used the elitist genetic algorithm and proposed an improved neighbour path initialization method to solve the EV routing problem [24]. Hien *et al.* proposed a greedy search algorithm GSGA inspired by clustering [25]. Wang *et al.* used Montecarlo tree search algorithm to improve the genetic algorithm, taking into account the flexibility of the paths and saving allocation time. The more complex the composition of the road network, the better the results obtained by the algorithm [26]. In 2023, Amiri *et al.* proposed developed two meta-heuristic algorithms including Non-dominated Sorting Genetic Algorithm II (NSGA-II) and Adaptive Large Neighbourhood Search (ALNS), and combined with multi-objective solution methods (e.g. weighted and epsilon constraints and hybrid methods) [27].

In this paper, based on the previous research, firstly, established an integer hybrid planning model for the EVRPTW problem that comprehensively considers multiple factors, including vehicle fixed cost, energy cost, time cost, and penalty cost for violating constraints, which is closer to the actual engineering background. Meanwhile, the Time of use Pricing and Area price are introduced, and the EVs into the grid based on the V2G model realize the formation of a smart microgrid, which effectively acquires and delivers energy by means of the EVs swap under the dual perspectives of time and space. In particular, the use of the A* algorithm in the initializing population step allows for a better initial solution, which avoids the problem of slow iteration speed and satisfies the stochastic nature of genetic algorithms. Experiments demonstrate that the GA-A* algorithm performs well in solving the EVRPTW problem, applies to different sizes of arithmetic cases, and outperforms the traditional GA algorithm and other path planning algorithms. In addition, it helps the smart microgrid not only achieves the successful reduction of the system's load peak-to-valley difference rate and mean square deviation based on the V2G model by incorporating EVs into the grid, effectively utilizing EVs to reduce the grid's volatility and improve the quality of power but also enables the EVs to rationally distribute energy in the grid by prioritizing the delivery of energy during peak hours of power consumption and acquiring energy during the trough hours of power consumption, which achieves a reasonable distribution of power.

2. Methodology

2.1 Math model

The main problem investigated in this paper is the Electric-Vehicle Routing Problem with Time Windows (EVRPTW), which is described as a number of EVs transporting the required supplies from a distribution center to various customer nodes, where they can be charged or discharged at a charging station, which adopts battery swapping. On the one hand, swapping batteries simplifies the model complexity, on the other hand, the replaced batteries can be used as part of a microgrid for charging and discharging operations, bringing profit to the grid operator. Verma confirmed that battery swapping leads to better path planning options for electric vehicles [28]. The objective function is to minimize the total cost including vehicle fixed cost, energy cost which considers the temperature coefficient, time cost, penalty cost for violating the load constraint, penalty cost for violating the time window, penalty cost for violating the power constraint, and charging and discharging cost.

EVs are subject to constraints:

- Each customer has a fixed amount of demand and is allowed to be served by only one vehicle and only once;
- EVs exist that the maximum load cannot exceed the maximum load capacity;
- EVs have a battery capacity that cannot exceed the maximum battery capacity or a battery capacity that is less than zero;

- EVs depart from the distribution center in a fully charged state and eventually return to the charging center;
- EVs are allowed to visit the charging station multiple times;
- EVs are required to arrive by the latest time window or pay a penalty cost;
- EVs arriving early at the customer node are not required to pay penalty costs, only time costs according to time.

In the problem, the Time of use Pricing and area price are introduced to charge different prices for 24 hours of the day according to different periods and in different areas. Higher prices are charged during the periods and areas with excessive electricity demand; conversely, the prices are lowered. The base load is referred to the data provided by Xiao *et al.* [29]. Fig. 1 shows the 24 hours area price and the grid load at each moment, and the area price is set with reference to the data provided by Fu *et al.* [30]. Fig. 2 shows the area price and the grid load in different areas. Huo has demonstrated that ambient temperature has a great influence on the energy consumption of EVs, and Fig. 3 shows the temperature curve fitted according to the data of this article [31]. Fig. 4 shows the annual time-averaged temperature profile for Beijing in 2022, with data from the National Oceanic and Atmospheric Administration (NOAA) National Center for Environmental Information (NCEI).

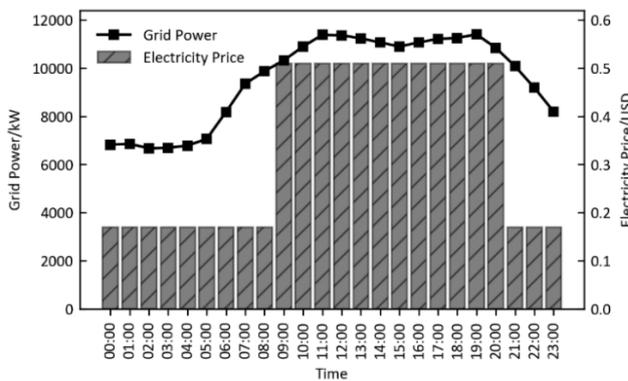


Fig. 1 Time of use pricing and grid loads

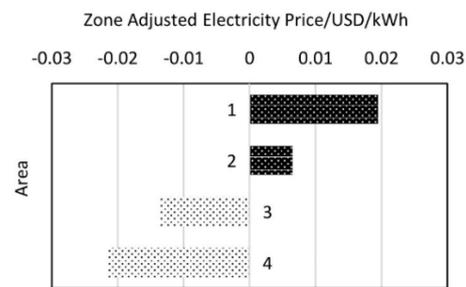


Fig. 2 Area price adjustment

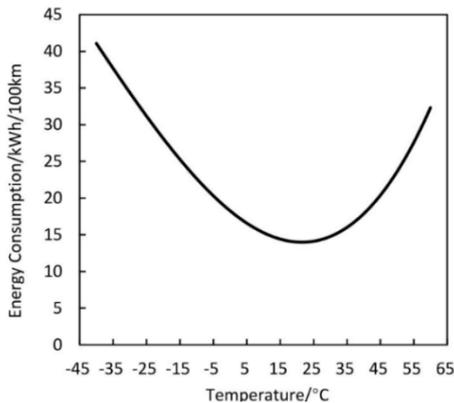


Fig. 3 Temperature-energy consumption per 100 km curve

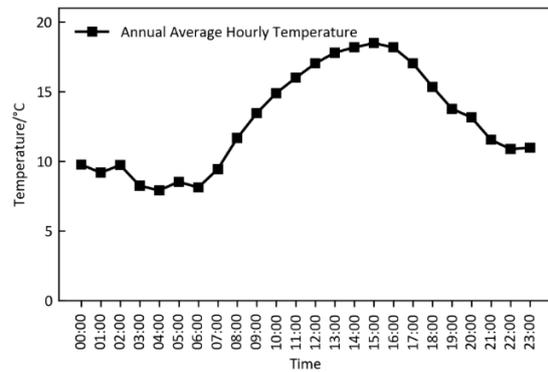


Fig. 4 Beijing's annual average temperature time curve

$$y = 0.000052307x^3 + 0.0069952x^2 - 0.37395x + 18.276 \tag{1}$$

Eq. 1 shows the fitted temperature-energy consumption per 100 km curve.

Due to the short driving range of EVs, they need to be recharged at the charging station during the driving process. In this paper, electric energy is recharged by swapping batteries and connected to the power grid to deliver/acquire energy. According to the impact of the Time of use Pricing and the area price, electricity is sold to the power grid for profit. V2G technology is introduced by simulating V2G based on the Time of use Pricing and the area price. This paper divides a day into 24 time periods, each with different electricity prices. During periods of high electricity demand, it is the peak period of electricity consumption, and the Time of use Pricing is

higher. Conversely, it is a low valley stage, and the Time of use Pricing is lower. Divide the entire map into different regions based on different datasets. In areas of peak electricity consumption, area prices are higher; conversely, area prices are lower. When the vehicle is charging, replace the battery with a fully charged battery at the charging station; When the vehicle is discharged, replace the battery at the charging station with enough power to support the next charging station or distribution center.

The EVRPTW model studied in this paper is, given a graph $G = (N \cup E, A)$, where the point set $N = \{0, 1, \dots, n\}$ represents a set of n customers' points, where 0 represents a distribution center, $E = \{n, n + 1, \dots, n + m\}$ represents m charging stations, $A = \{(i, j) | i, j \in N \cup E, i \neq j\}$ represents all connected arcs in $N \cup E$.

The symbol descriptions in the model are shown in Table 1.

Table 1 Symbol description

Symbol	Instructions
O	Distribution center
N	Customer node set
K	Set of the number of EVs used
E	Charging Station set
V	All node set
d_{ij}	Node i to node j traveling distance
D_i	Customer node i demand
L_{mc}	Max. loading capacity of the vehicle
L_{ik}	Remaining load at arrival of vehicle k at node i
B_Q	Max. battery capacity of vehicle
B_{aik}	Remaining capacity at arrival of vehicle k at node i
B_{lik}	Remaining capacity at left of vehicle k at node i
E_i	Customer node i 's earliest service time, $i \in N$
L_i	Customer node i 's latest service time, $i \in N$
t_{aik}	Time at arrival of vehicle k at node i , $i \in N$
t_{lik}	Time at leaving of vehicle k at node i , $i \in N$
t_{wik}	Waiting Time of vehicle k at node i , $i \in N$
t_{mik}	Missing Time of vehicle k at node i , $i \in N$
t_{sik}	Service/Swap Time of vehicle k at node i , $i \in N$
t_{ijk}	Time at driving of vehicle k from node i to node j , $i, j \in N$
v	velocity of vehicles traveling in distribution
μ	Battery consumption rate
$Tem_{\gamma k}$	γ time vehicle k traveling ambient temperature coefficient
C_d	Cost per distance
C_1	Penalty costs for violating load constraints
C_2	Penalty costs for violating the time window
C_3	Penalty costs for violating electricity constraints
C_t	Cost per time
$S_{\gamma\theta}$	γ time area θ price
$P_{L\gamma\theta}$	No EVs in the original grid γ time area θ power

The EVRPTW model proposed in this paper is improved from the mathematical model proposed by Li *et al.* [32], and on this basis, increases the Time of use Pricing, Area price, and ambient temperature coefficient. On the one hand, it is closer to the real working environment, and on the other hand, it can help regulate and store electricity in space, which can better improve energy quality and play a role in energy distribution. According to the description of the EVRPTW problem, the objective function includes vehicle fixed cost, energy cost which considers the temperature coefficient, time cost, penalty cost for violating load constraints, penalty cost for violating time windows, penalty cost for violating energy constraints, and charging and discharging cost. A mathematical model for this problem is established.

Decision variables:

$$X_{ijk} = \begin{cases} 1, & \text{Vehicle } k \text{ from customer } i \text{ to customer } j \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned}
 Y_k &= \begin{cases} 1, & \text{vehicle } k \text{ violated the power constraints} \\ 0, & \text{otherwise} \end{cases} \\
 \min C &= C_0 \sum_{k \in K} X_k + C_1 \sum_{k \in K} \max[0, D_k - L_{mc}] + C_2 \sum_{i \in V} \sum_{k \in K} t_{mik} \\
 &+ C_3 \sum_{k \in K} Y_k + C_d \sum_{i \in V} \sum_{j \in V} \sum_{k \in K} \sum_{\gamma=1}^{24} Tem_{\gamma k} d_{ij} X_{ijk} \tag{2} \\
 &+ C_t \sum_{i \in V} \sum_{j \in V} \sum_{k \in K} (t_{ijk} + t_{sjk}) + \sum_{\gamma=1}^{24} \sum_{\theta=1}^n \sum_{k \in K} |P_{\gamma k}| S_{\gamma \theta} \\
 &\sum_{i \in N, i \neq j} X_{ijk} = 1, \forall j \in N, k \in K \tag{3} \\
 &\sum_{j \in V, i \neq j} X_{0jk} \leq 1, k \in K \tag{4} \\
 &\sum_{i \in V, i \neq j} X_{ijk} - \sum_{i \in V, i \neq j} X_{jik} = 0, \forall j \in V, k \in K \tag{5} \\
 &0 \leq D_j \leq L_{ik} - D_i, i \in V, j \in V, k \in K \tag{6} \\
 &0 \leq \sum_{i \in N} D_{ik} \leq L_{mc}, \forall k \in K \tag{7} \\
 &0 \leq B_{ajk} \leq B_{lik} + Tem_{\gamma k} d_{ij} X_{ijk} \mu, \forall i \in N, j \in N, k \in K \tag{8} \\
 &\sum_{i, j \in V} Tem_{\gamma k} d_{ij} X_{ijk} \mu \leq B_{li} \leq B_Q, i \in G \tag{9} \\
 &0 \leq B_{aik} \leq B_Q, \forall i \in V, k \in K \tag{10} \\
 &B_{aik} = B_{lik}, \forall i \in V, k \in K \tag{11} \\
 &t_{wik} = \max[0, e_i - t_{aik}], \forall i \in N, k \in K \tag{12} \\
 &t_{mik} = \max[0, t_{aik} - l_i], \forall i \in N, k \in K \tag{13} \\
 &t_{lik} = t_{aik} + t_{sik} + t_{wik}, \forall i \in N \cup G \tag{14} \\
 &t_{ijk} = \frac{d_{ijk}}{v}, \forall i \in V, j \in V, k \in K \tag{15} \\
 &t_{ajk} = (t_{lik} + t_{ijk}) X_{ijk}, \forall i \in V, j \in N \cup G, k \in K \tag{16} \\
 &X_{ijk} = \{0, 1\}, \forall i \in V, j \in V, k \in K \tag{17} \\
 &Y_k = \{0, 1\}, \forall k \in K \tag{18}
 \end{aligned}$$

The objective function (Eq. 2) represents the minimum total cost, including vehicle fixed cost, energy consumption cost, time cost, charging and discharging cost, and penalty cost. Among them $S_{\gamma\theta}$ are the Time of use Pricing and Area price. When the battery is discharged, $S_{\gamma\theta}$ is negative, i.e. when the cost of charging and discharging is negative, it means that the charging and discharging process is generally profitable. Constraint Eq. 3 indicates that each customer is served only once. In the cost of time penalty, waiting time is not included in the penalty, but rather placed within the cost of time. We believe that this can improve the efficiency of resource allocation. Constraint Eq. 4 indicates that only one transport vehicle is arranged for each distribution route. Constraint Eq. 5 indicates equal number of vehicles entering and exiting. Con-

straint Eq. 6 indicates that the demand of the vehicle at the next node cannot be higher than the remaining demand of the vehicle at the current node minus the demand at the current node. Constraint Eq. 7 indicates that all demands of a transportation route must not exceed the maximum load capacity. Constraint Eq. 8 indicates that the remaining power upon reaching the next node must not be less than the current node's power minus the power consumption from that node to the next node. Constraint Eq. 9 indicates that the energy consumption after charging and discharging at the charging station is not less than the energy consumption of the sub path and not higher than the maximum energy consumption. Constraint Eq. 10 indicates that the remaining electricity at any point is not negative. Constraint Eq. 11 indicates that the remaining electricity is unchanged before and after the vehicle visits the customer node. Constraint Eq. 12 represents the calculation method of waiting time. Constraint Eq. 13 represents the calculation method of late time. Constraint Eq. 14 indicates that the time of a vehicle leaves the node is the sum of arrival time, waiting time, and service time. Constraint Eq. 15 represents the calculation method of the time taken by the vehicle from node i to node j . Constraint 16 indicates that the time for the vehicle to reach the next node is equal to the sum of the time to leave the current node and the time spent on the journey. Constraints Eqs. 17 and 18 are binary 0-1 variables.

2.2 Algorithm

To solve the Electric-Vehicle Routing Problem with Time Windows, this paper proposes an improved genetic algorithm GA-A*. Genetic algorithm has the advantages of starting from the population, having potential for parallelism, using evaluation functions for inspiration, simple processes, probability mechanisms for iteration, and randomness. However, genetic algorithms have certain dependencies in initial population selection, which may lead to the problem of falling into local optimal solutions.

To overcome these problems, this paper introduces the A* algorithm as part of the improved genetic algorithm. The A* algorithm is a heuristic search algorithm that guides the search process by evaluating the cost function and heuristic function of nodes. It has high efficiency and accuracy in path search. The purpose of introducing the A* algorithm is to optimize the initial set of points by solving for the optimal path and cutting it into seed nodes. Then, by randomly creating other nodes as genes, the initialized population is formed. Such an improvement measure aims to improve the search efficiency and effectively avoid the occurrence of local optimal solutions.

This paper chooses to introduce the A* algorithm. Firstly, it can evaluate nodes through heuristic functions to quickly find the optimal path. Secondly, the A* algorithm has good performance and accuracy in solving the optimal path problem. By applying the A* algorithm to the optimization of the initial point set, the quality of the initial population can be improved, providing a better starting point for the search process of genetic algorithms. Finally, this improvement by introducing the A* algorithm helps to improve the overall algorithm's ability to solve the EVRPTW and provides better results.

The pseudo-code of the GA-A* algorithm is shown in Fig. 5. Through this improvement, this paper aims to overcome the limitations of traditional genetic algorithms and improve the efficiency and quality of solving path planning problems.

Firstly, based on the A* algorithm, the optimal path passing through all nodes is obtained and the path is divided into sub-paths according to the needs of the customer nodes. Then sort the remaining unincorporated nodes by time window order, and insert randomly signed charging station nodes to generate multiple genes as the initial population. Using decimal signed encoding to represent genes, including information such as customer nodes, charging nodes, and distribution centers. The fitness function considers vehicle costs, energy consumption costs, time costs, and penalty costs for constraint violations. Selection, using the roulette selection method. Crossover, using the first 1/3 encoding after crossover to replace the last 1/3 encoding before crossover. Variation, using an elitist strategy to retain elitist individuals, varies general genes, generates multiple genes, and calculates the fitness to retain the highest term. Multiple iterations are performed to obtain the final path encoding.

Algorithm 1 GA-A* algorithm

Input: $V, A, d_{ij}, Station_area, D_i, t_{sik}, t_{ijk}, E_i, L_i, C_0, C_1, C_2, C_3, C_d, C_t$
Output: *Best path*

```

class S object:
1: Dis  $\leftarrow S.var$ 
2: npath  $\leftarrow Priority\_Que()$ 
3: npath.S
4: while npath  $\neq \emptyset$  do
5:   if  $len(npath) = num\_cities$  then
6:     path  $\leftarrow npath$ 
7:   end if
8:   visited[current]  $\leftarrow True$ 
9:   for neighbor  $\leftarrow 0$  to  $num\_cities - 1$  do
10:    if neighbor in charging_stations.keys() then
11:      continue
12:    end if
13:    if not visited[neighbor] then
14:      has_gone_to_charging  $\leftarrow 0$ 
15:      if has_gone_to_charging = 0 then
16:        has_gone_to_charging  $\leftarrow 1$ 
17:        find min cost with Station
18:      else if has_gone_to_charging == 1 then
19:        find min cost without Station
20:      end if
21:      continue
22:    end if
23:  end for
24: end while
25: path  $\leftarrow npath$ 
26: for index, pos in enumerate(path) do
27:   Path encoding is segmented based on demand
28: end for
29: for sub_path in (path) do
30:   Each sub_path code arranged in TW and randomly sorted by grouping the
   remaining codes
31: end for
32: for i  $\leftarrow 1$  to k do
33:   chosen_pop  $\leftarrow choose(pop)$ 
34:   crossed_pop  $\leftarrow cross(chosen\_pop)$ 
35:   pop  $\leftarrow mergeGenes(pop, crossed\_pop)$ 
36:   pop  $\leftarrow vary(pop)$ 
37:   key  $\leftarrow gene : fitness\_pop(gene)[0]$ 
38:   pop.sort(reverse = True, key = key)
39: end for

```

Fig. 5 GA-A* pseudo-code

2.3 Experimental design

To evaluate the performance of the model and algorithm. In this paper, 16 test cases were selected. Test cases are derived from a shared database [33]. In these cases, each customer node has different requirements, and each example has a different charging station and number of customers. For example, the case "R-2-C-40" means that there are two charging stations and forty customers in the case. To test the performance of the GA-A* algorithm, experiments were designed to test it against the conventional GA, and referred to the data made by Li et al using adaptive large neighbourhood search (ALNS), large neighbourhood search (LNS), and variable neighbourhood search (VNS) algorithms [32] for comparison. The algorithm is written in Python and the test platform is Intel Xeon E5-2680 v4.

3. Result

3.1 Simulation results

The algorithm was used to test the optimal path under 16 groups of different cases, Some of the initial maps are shown in Fig. 6 for 30, 50, 70 and 120 charging station nodes.

Fig. 7 shows the paths plotted from the optimal paths derived from the GA-A* operation with other methods.

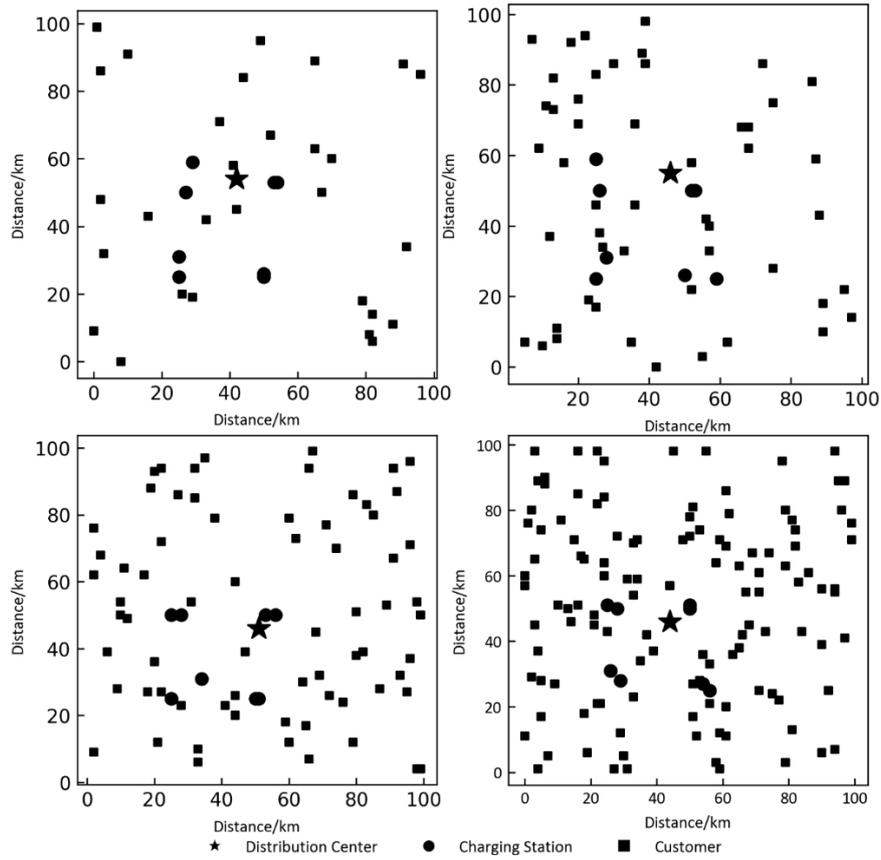


Fig. 6 Initial map

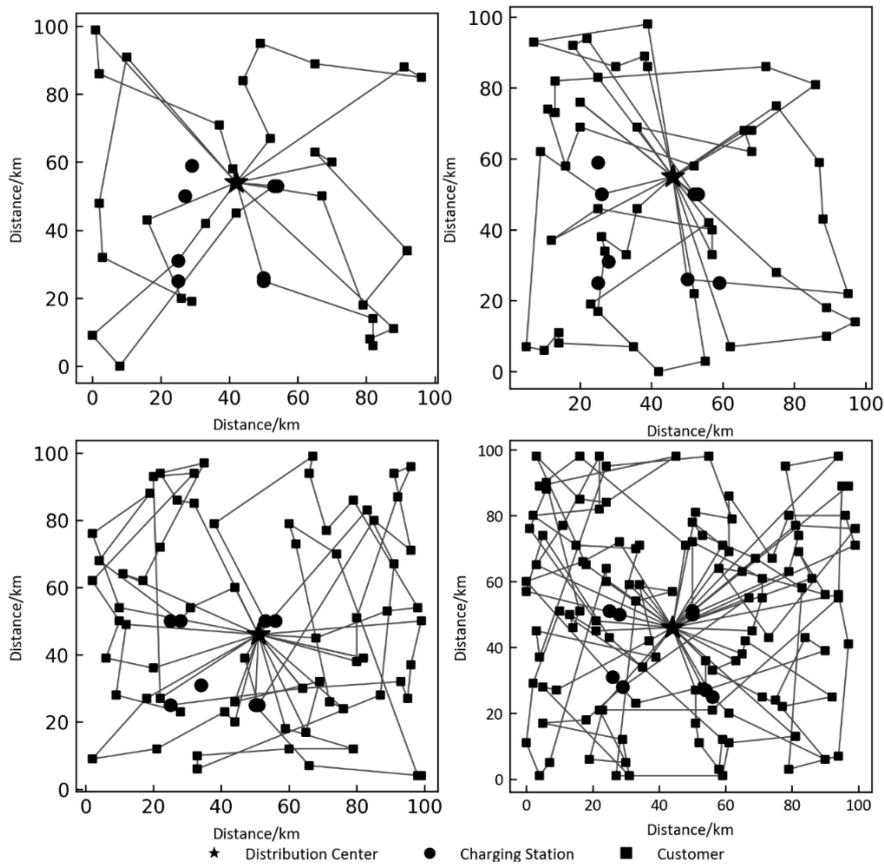


Fig. 7 Route map

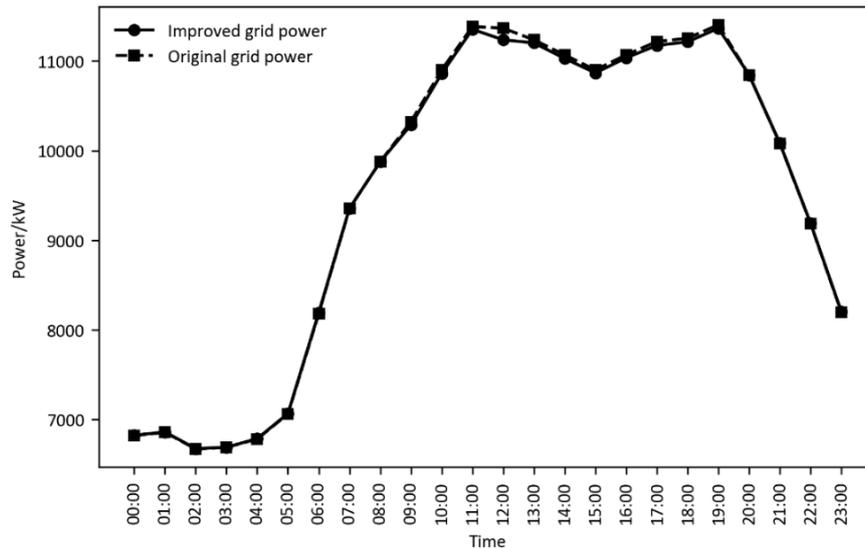


Fig. 8 Grid load comparison

As shown in Fig. 7, the algorithm can find the route that fulfils the conditions and ensures the correctness of the solution. After that, the paper tests the effect of the GA-A* algorithm on the power grid at different times, and the experimental results show that the GA-A* algorithm has a good effect on power sharing and power quality.

Fig. 8 shows better power distribution under the GA-A* algorithm scheduling. Most of the charging time is concentrated between 21:00 and 3:00 the next day, which is in the off-peak period of electricity consumption. The discharge time is concentrated from 9:00 to 19:00, which is in the peak period of power consumption, the peak load decreases from the original 11403.73 kW to 11365.10 kW, and the peak differential rate decreases from the original 41.50-41.27 %. Charging when the electricity price is low and discharging when the electricity price is high, fully shows that the scheduling function of the GA-A* algorithm can cut peak and fill the valley and reduce the charging cost. As shown in Table 2.

Due to the existence of area price, the area price can be used to regulate loads in different areas of the grid. As shown in Fig. 9, if the charging station is in a region with low electricity prices, the route is more concerned with charging in this region; if it is in a region with high electricity prices, the route is more concerned with discharging in this region.

Table 2 Load characteristic index in the different scenarios

scenarios	Max. load (kW)	Min. load (kW)	mean square	peak-to-valley ratio (%)
original grid	11403.73	6670.81	1846.76	41.50
GA-A*	11365.10	6675.15	1825.69	41.27

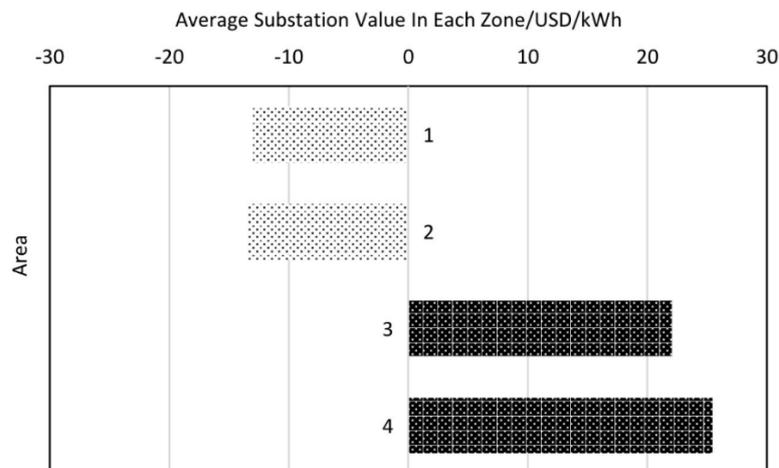


Fig. 9 Zonal grid load changes

3.2 Comparison with other methods

Fig. 10 shows the comparison between the GA-A* algorithm and the traditional GA algorithm. As shown in the figure, the iteration speed of GA-A* is significantly faster than that of the GA algorithm, and it enters the low value region 89 times and completes the iteration quickly. The shortest distance produced by GA-A* is also significantly smaller than the optimal solution produced by GA.

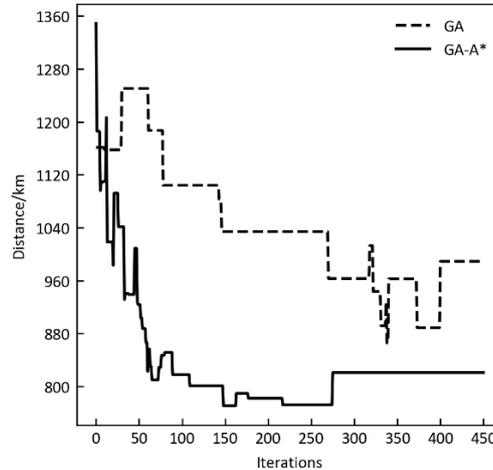


Fig. 10 GA-A* versus GA

Table 3 shows the comparison results between GA-A* and the other three algorithms. It is found that the GA-A* algorithm has a better optimization effect than the other three algorithms. In comparison to LNS, VNS, and ALNS, the average reduction is 281.41, -67.43, and 207.80 km, respectively. The results show that the optimization effect of GA-A* is significantly higher than that of LNS and ALNS.

Table 3 Load characteristic index in the different scenarios

case	optimal solution /km			
	GA-A*	LNS	VNS	ALNS
r-8-c-120	3282.46	3617.30	3160.50	3602.10
r-8-c-70	2067.05	2382.70	1938.40	2242.20
r-8-c-50	1401.98	1672.30	1419.90	1646.70
r-8-c-30	947.25	1333.30	983.60	1145.50
r-6-c-120	2428.78	2132.90	1841.20	2479.10
r-6-c-70	1791.31	2196.80	1767.40	2057.80
r-6-c-50	1536.47	2417.70	1556.90	1784.40
r-6-c-30	820.92	1469.90	828.50	1009.10
r-4-c-80	2249.17	2369.50	2181.10	2465.90
r-4-c-70	1975.49	2305.20	2117.10	2310.90
r-4-c-50	1469.40	1532.80	1327.90	1583.60
r-4-c-30	849.50	1087.90	881.10	1043.10
r-2-c-90	2455.17	2722.80	2285.50	2707.10
r-2-c-70	1934.04	2054.20	1842.20	2099.90
r-2-c-50	1369.70	1575.50	1357.60	1569.80
r-2-c-30	873.24	1083.70	884.20	1029.50

4. Discussion

In the process of solving the EVRPTW problem, in order to reduce the total cost in the transportation process and make the model closer to the actual engineering background, we comprehensively consider several factors and establish the model by using integer mixed programming. These factors include the fixed cost of the vehicle, energy cost which considers the temperature coefficient, the cost of time, the penalty cost of violating the load constraint, the penalty cost of violating the time window, and the penalty cost of violating the power constraint and the charge

and discharge cost. Through such comprehensive consideration, we obtained a model of EVRPTW that is more in line with the actual engineering context.

By introducing the Time of use Pricing and Area price, we have reached the V2G model, integrating EVs into the grid, and realizing the transportation and exchange of energy is realized. By means of power swapping for electric vehicles, we have achieved the purpose of forming a smart microgrid. In this mode, EVs can choose to deliver energy during peak hours of electricity consumption and obtain energy during low hours of electricity consumption according to the demand of the power grid, rationalizing the distribution of electric energy, reducing the peaks and valleys and volatility of the load on the power grid, and improving the quality of electric energy. V2G mode makes EVs become part of the power grid, participate in the transportation and exchange of energy, and make full use of the energy storage capacity and flexibility of EVs through intelligent energy scheduling and control.

In terms of algorithm, we improve the steps of initializing the population and introduce the A* algorithm to obtain a better initial solution. By combining GA and A* algorithm, we not only avoid the problem of slow iteration speed of the genetic algorithm but also meet the randomness requirement of GA. Experimental results show that this GA-A* algorithm can provide correct path planning, and has better iteration speed and optimization quality when solving EVRPTW problem. Compared with the traditional genetic algorithm, our algorithm shows obvious advantages in iteration speed and optimal solution quality. Compared with other different types of path planning algorithms, such as LNS, ALNS, and VNS algorithms, our algorithm also shows better performance and can get good results in different scale examples. In some specific cases as shown in Table 3 it can be seen that the distance of GA-A* is less than the solution derived by VNS for small scale cases with eighty client nodes, but in large scale cases VNS shows better performance. This paper argues that the implementation of the VNS algorithm searches a wider space of solutions in variable neighbourhoods, no matter the size of the arithmetic case can find a better solution by means of global search. VNS follows this feature for solutions at different scales, whereas GA-A*, which has fixity in its initialisation, shows better performance in small scale solving. In this paper, we argue that GA-A* has significantly outperformed other algorithms under the path length perspective for small-scale problems, and some algorithms for large-scale problems; it outperforms other algorithms under the perspective for spatio-temporal deployment of energy consumption. The algorithm has been initialised in such a way as to find the optimal solution passing through all nodes, so that there is a certain degree of similarity between the initialised population and the final solution for the small-scale example in the global solution; the algorithm outperforms other algorithms in terms of energy consumption because it takes into account the spatio-temporal tariffs. There are also some weaknesses in that the initialised population is still somewhat fixed even with the addition of random values, so larger than large-scale arithmetic cases lack global search capability. So introduction of the A* algorithm as part of the initialization allows the genetic algorithm to start searching for a better initial solution. This combination can fully utilize the advantages of the two algorithms and improve the search efficiency and solution quality of the algorithm.

In summary, by comprehensively considering multiple practical factors and establishing an integer mixed programming model, introducing Time of use Pricing and Area price as well as the V2G model, and combining the genetic algorithm and A* algorithm for optimization, we have made remarkable progress in solving EVRPTW problem. Our model and algorithm can better conform to the actual engineering background, obtain a better solution, improve the efficiency and quality of power grid energy management, and bring a positive impact on the operation of smart microgrid and the entire power grid. Such as Wang *et al.* used heuristics and collision avoidance algorithms for collaborative scheduling planning of multiple AGVs [34]. GA-A* can be generalised to AGVs in industrial manufacturing for logistics optimisation tasks such as optimal routing through improvements.

5. Conclusion

To solve the EVRPTW, this paper comprehensively considers multiple factors in reality and takes the minimum total cost as the model proposes the GA-A* algorithm which introduces the A* algorithm, and tests 16 cases of four groups of different scales. Experiments show that compared with the traditional GA algorithm and other algorithms, it has a certain optimization effect. The algorithm itself can provide a solution to the np-hard problem and get the right route. At the same time, based on the V2G mode, the smart microgrid can reduce fluctuations and improve power quality, and the reasonable spatio-temporal allocation of EVs plays a role in peaking and valley filling for the entire power grid.

The GA-A* algorithm introduced in this paper can provide better path planning in the search process, but it may face the problem of too long computation time when dealing with large-scale problems. Therefore, for larger scale EVRPTW problems, it may be necessary to further optimize the efficiency of the algorithm. At the same time, in practical problems, there are often multiple conflicting objectives that need to be optimized, The GA-A* algorithm can be further extended. Therefore, to solve the above problems, we can consider further optimizing the efficiency of the algorithm, such as introducing a pruning strategy or parallel computing. Multi-Objective Optimization can be supported by improvement, such as considering the problem of simultaneously optimizing total cost and total time.

References

- [1] Yilmaz, M., Krein, P.T. (2013). Review of the impact of vehicle-to-grid technologies on distribution systems and utility interfaces, *IEEE Transactions on Power Electronics*, Vol. 28, No. 12, 5673-5689, doi: [10.1109/TPEL.2012.2227500](https://doi.org/10.1109/TPEL.2012.2227500).
- [2] Abdulaal, A., Cintuglu, M.H., Asfour, S., Mohammed, O.A. (2017). Solving the multivariant EV routing problem incorporating V2G and G2V options, *IEEE Transactions on Transportation Electrification*, Vol. 3, No. 1, 238-248, doi: [10.1109/TTE.2016.2614385](https://doi.org/10.1109/TTE.2016.2614385).
- [3] Norouzi, F., Hoppe, T., Ramirez Elizondo, L., Bauer, P. (2022). A review of socio-technical barriers to Smart Microgrid development, *Renewable and Sustainable Energy Reviews*, Vol. 167, Article No. 112674, doi: [10.1016/j.rser.2022.112674](https://doi.org/10.1016/j.rser.2022.112674).
- [4] Yoldaş, Y., Önen, A., Muyeen, S.M., Vasilakos, A.V., Alan, İ. (2017). Enhancing smart grid with microgrids: Challenges and opportunities, *Renewable and Sustainable Energy Reviews*, Vol. 72, 205-214, doi: [10.1016/j.rser.2017.01.064](https://doi.org/10.1016/j.rser.2017.01.064).
- [5] Khan, A.A., Naeem, M., Iqbal, M., Qaisar, S., Anpalagan, A. (2016). A compendium of optimization objectives, constraints, tools and algorithms for energy management in microgrids, *Renewable and Sustainable Energy Reviews*, Vol. 58, 1664-1683, doi: [10.1016/j.rser.2015.12.259](https://doi.org/10.1016/j.rser.2015.12.259).
- [6] Keskin, M., Çatay, B. (2016). Partial recharge strategies for the electric vehicle routing problem with time windows, *Transportation Research Part C: Emerging Technologies*, Vol. 65, 111-127, doi: [10.1016/j.trc.2016.01.013](https://doi.org/10.1016/j.trc.2016.01.013).
- [7] Keskin, M., Çatay, B. (2018). A matheuristic method for the electric vehicle routing problem with time windows and fast chargers, *Computers & Operations Research*, Vol. 100, 172-188, doi: [10.1016/j.cor.2018.06.019](https://doi.org/10.1016/j.cor.2018.06.019).
- [8] Wang, D., Zhou, H., Feng, R. (2019). A two-echelon vehicle routing problem involving electric vehicles with time windows, *Journal of Physics: Conference Series*, Vol. 1324, No. 1, Article No. 012071, doi: [10.1088/1742-6596/1324/1/012071](https://doi.org/10.1088/1742-6596/1324/1/012071).
- [9] 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).
- [10] Raeesi, R., Zografos, K.G. (2020). The electric vehicle routing problem with time windows and synchronised mobile battery swapping, *Transportation Research Part B: Methodological*, Vol. 140, 101-129, doi: [10.1016/j.trb.2020.06.012](https://doi.org/10.1016/j.trb.2020.06.012).
- [11] Zhu, Y., Lee, K.Y., Wang, Y. (2021). Adaptive elitist genetic algorithm with improved neighbor routing initialization for electric vehicle routing problems, *IEEE Access*, Vol. 9, 16661-16671, doi: [10.1109/ACCESS.2021.3053285](https://doi.org/10.1109/ACCESS.2021.3053285).
- [12] Deng, J., Li, J., Li, C., Han, Y., Liu, Q., Niu, B., Liu, L., Zhang, B. (2021). A hybrid algorithm for electric vehicle routing problem with nonlinear charging, *Journal of Intelligent & Fuzzy Systems*, Vol. 40, No. 3, 5383-5402, doi: [10.3233/JIFS-202164](https://doi.org/10.3233/JIFS-202164).
- [13] Bac, U., Erdem, M. (2021). Optimization of electric vehicle recharge schedule and routing problem with time windows and partial recharge: A comparative study for an urban logistics fleet, *Sustainable Cities and Society*, Vol. 70, Article No. 102883, doi: [10.1016/j.scs.2021.102883](https://doi.org/10.1016/j.scs.2021.102883).
- [14] Lin, B., Ghaddar, B., Nathwani, J. (2021). Electric vehicle routing with charging/discharging under time-variant electricity prices, *Transportation Research Part C: Emerging Technologies*, Vol. 130, Article No. 103285, doi: [10.1016/j.trc.2021.103285](https://doi.org/10.1016/j.trc.2021.103285).

- [15] Lin, B., Ghaddar, B., Nathwani, J. (2022). Deep reinforcement learning for the electric vehicle routing problem with time windows, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 8, 11528-11538, doi: [10.1109/TITS.2021.3105232](https://doi.org/10.1109/TITS.2021.3105232).
- [16] Erdelić, T., Carić, T. (2022). Goods delivery with electric vehicles: Electric vehicle routing optimization with time windows and partial or full recharge, *Energies*, Vol. 15, No. 1, Article No. 285, doi: [10.3390/en15010285](https://doi.org/10.3390/en15010285).
- [17] Niu, X.Y., Liu, S.F., Huang, Q.L. (2022). End-of-line delivery vehicle routing optimization based on large-scale neighbourhood search algorithms considering customer-consumer delivery location preferences, *Advances in Production Engineering & Management*, Vol. 17, No. 4, 439-454, doi: [10.14743/apem2022.4.447](https://doi.org/10.14743/apem2022.4.447).
- [18] Liu, Q., Xu, P., Wu, Y., Shen, T. (2022). A hybrid genetic algorithm for the electric vehicle routing problem with time windows, *Control Theory and Technology*, Vol. 20, No. 2, 279-286, doi: [10.1007/s11768-022-00091-1](https://doi.org/10.1007/s11768-022-00091-1).
- [19] Ding, N., Yang, J., Han, Z., Hao, J. (2022). Electric-vehicle routing planning based on the law of electric energy consumption, *Mathematics*, Vol. 10, No. 17, Article No. 3099, doi: [10.3390/math10173099](https://doi.org/10.3390/math10173099).
- [20] Kumar, A., Aggarwal, A., Rani, A., Bedi, J., Kumar, R. (2023). A mat-heuristics approach for electric vehicle route optimization under multiple recharging options and time-of-use energy prices, *Concurrency and Computation: Practice and Experience*, Vol. 35, No. 27, Article No. e7854, doi: [10.1002/cpe.7854](https://doi.org/10.1002/cpe.7854).
- [21] Kempton, W., Tomić, J. (2005). Vehicle-to-grid power fundamentals: Calculating capacity and net revenue, *Journal of Power Sources*, Vol. 144, No. 1, 268-279, doi: [10.1016/j.jpowsour.2004.12.025](https://doi.org/10.1016/j.jpowsour.2004.12.025).
- [22] Shao, S., Guan, W., Bi, J. (2018). Electric vehicle-routing problem with charging demands and energy consumption, *Intelligent Transport System*, Vol. 12, No. 3, 202-212, doi: [10.1049/iet-its.2017.0008](https://doi.org/10.1049/iet-its.2017.0008).
- [23] Wang, X.P., Wang, M., Ruan, J.H., Li, Y. (2018). Multi-objective optimization for delivering perishable products with mixed time windows, *Advances in Production Engineering & Management*, Vol. 13, No. 3, 321-332, doi: [10.14743/apem2018.3.293](https://doi.org/10.14743/apem2018.3.293).
- [24] Zhu, Y., Li, C., Lee, K.Y. (2022). The NR-EGA for the EVRP problem with the electric energy consumption model, *Energies*, Vol. 15, No. 10, Article No. 3681, doi: [10.3390/en15103681](https://doi.org/10.3390/en15103681).
- [25] Hien, V.Q., Dao, T.C., Binh, H.T.T. (2023). A greedy search based evolutionary algorithm for electric vehicle routing problem, *Applied Intelligence*, Vol. 53, No. 3, 2908-2922, doi: [10.1007/s10489-022-03555-8](https://doi.org/10.1007/s10489-022-03555-8).
- [26] Wang, Y.D., Lu, X.C., Song, Y.M., Feng, Y., Shen, J.R. (2022). Monte Carlo Tree Search improved Genetic Algorithm for unmanned vehicle routing problem with path flexibility, *Advances in Production Engineering & Management*, Vol. 17, No. 4, 425-438, doi: [10.14743/apem2022.4.446](https://doi.org/10.14743/apem2022.4.446).
- [27] Amiri, A., Zolfagharinia, H., Amin, S.H. (2023). A robust multi-objective routing problem for heavy-duty electric trucks with uncertain energy consumption, *Computers & Industrial Engineering*, Vol. 178, Article No. 109108, doi: [10.1016/j.cie.2023.109108](https://doi.org/10.1016/j.cie.2023.109108).
- [28] Verma, A. (2018). Electric vehicle routing problem with time windows, recharging stations and battery swapping stations, *EURO Journal on Transportation and Logistics*, Vol. 7, No. 4, 415-451, doi: [10.1007/s13676-018-0136-9](https://doi.org/10.1007/s13676-018-0136-9).
- [29] Xiao, H., Pei, W., Kong, L. (2017). Multi-objective optimization scheduling method for active distribution network with large scale electric vehicles, *Transactions of China Electrotechnical Society*, Vol. 32, No. 202, 179-189.
- [30] Fu, R., Qiu, Y.S., Li, Y. (2007). A fuzzy clustering method for price partition based on the sensitivity of node price, *Power Demand Side Management*, No. 3, 19-22, doi: [10.3969/j.issn.1009-1831.2007.03.009](https://doi.org/10.3969/j.issn.1009-1831.2007.03.009).
- [31] Huo, Y.L. (2022). Research on energy consumption model energy-saving path planning of BEV, doi: [10.27162/d.cnki.gjlin.2022.004541](https://doi.org/10.27162/d.cnki.gjlin.2022.004541).
- [32] Li, J.L., Liu, Z.B., Wang, X.F. (2021). Distribution path optimization of electric vehicles considering charging and discharging strategy in smart grid, *Journal of South China University of Technology (Natural Science Edition)*, Vol. 49, No. 10, 31-40.
- [33] Zhang, S., Chen, M., Zhang, W., Zhuang, X. (2020). Fuzzy optimization model for electric vehicle routing problem with time windows and recharging stations, *Expert Systems with Applications*, Vol. 145, Article No. 113123, doi: [10.1016/j.eswa.2019.113123](https://doi.org/10.1016/j.eswa.2019.113123).
- [34] Wang, Y.J., Liu, X.Q., Leng, J.Y., Wang, J.J., Meng, Q.N., Zhou, M.J. (2022). Study on scheduling and path planning problems of multi-AGVs based on a heuristic algorithm in intelligent manufacturing workshop, *Advances in Production Engineering & Management*, Vol. 17, No. 4, 505-513, doi: [10.14743/apem2022.4.452](https://doi.org/10.14743/apem2022.4.452).

Optimizing rock breaking performance: The influence of chamfer on polycrystalline diamond compact (PDC) cutters

Ju, P.^{a,*}

^aCCTEG Xi'an Research Institute, Xi'an, P.R. China

ABSTRACT

Research on the rock-breaking performance of the Polycrystalline Diamond Compact (PDC) cutter has primarily focused on sharp cutters, often overlooking the influence of chamfer. Notably, the design of chamfer parameters has been largely unreported. In this study, we established a theoretical model of cutting force that takes chamfer into account. We analysed the primary and secondary relationships of four factors – back rake angle, depth of cut, chamfer angle, and chamfer length – on the force of the PDC cutter. This was done through a pseudo-level orthogonal level test. A numerical simulation, based on the Smooth Particle Hydrodynamic (SPH) method, was conducted to analyse the rock-breaking force and stress distribution characteristics of PDC cutters with different chamfer angles. Combined with a drop hammer impact test, we provided an optimized design of chamfer parameters. Our findings revealed that while the chamfer had a relatively minor influence on the force of the PDC cutter, it contributed to the optimal distribution of stress on the PDC cutter. This effectively protected the cutting edge and prevented early cracks and spalls of the cutter. When the chamfer angle was less than or equal to the back rake angle, the resultant force of the PDC cutter increased with the increase of the chamfer angle. However, when the chamfer angle was greater than the back rake angle, the resultant force of the PDC cutter first increased and then slightly decreased with the increase of the chamfer angle. Additionally, the resultant force of the PDC cutter increased approximately linearly with the increase of chamfer length. When the chamfer angle of the PDC cutter was between 30° and 45°, the fluctuation of the cutting force was relatively smooth, the rock-breaking process was stable, and the cutter's impact resistance energy was relatively higher. These findings will provide valuable guidelines for the design of chamfered PDC cutters.

ARTICLE INFO

Keywords:

Polycrystalline diamond compact (PDC) cutter;
Chamfer parameters;
Optimization;
Cutting force;
Theoretical analysis;
Numerical simulation;
Smooth Particle Hydrodynamic (SPH);
Stress characteristics

*Corresponding author:

jpnt2005@163.com
(Ju, P.)

Article history:

Received 25 October 2023
Revised 14 December 2023
Accepted 15 December 2023



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

1. Introduction

In the field of petroleum and coal geological drilling, PDC bit is widely used in soft to medium hard formations, thanks to its strong rock cutting ability and high drilling efficiency. As the direct rock breaking component of PDC bit, the performance of polycrystalline diamond compact (PDC) cutter determines the drilling effect and service life of PDC bit. The PDC cutter is synthesized by polycrystalline diamond layer and cemented carbide substrate at high temperature and high pressure environment, its shape is generally cylindrical, and it has the advantages of strong self-sharpening ability, good thermal stability and strong wear resistance. However, when drilling into hard and complex formation, PDC cutters will withstand serious impact and vibration, and the problems of cutter breakage and collapse occur frequently, which seriously restrict the performance of PDC bit.

In view of the above problems, a lot of studies have been carried out. Regarding the influence of back rake angle on the performance of PDC cutter, Vusal, Zhang *et al.* pointed out that the mechanical specific energy increases as back rake angle increases [1, 2]; Adzis *et al.* raised that the larger the back rake angle, the smaller the Von Mises stress induced in the cutter, and the cutter wear rate decreases as the back rake angle increases [3]. In view of the influence of depth of cut on the rock breaking force of PDC cutter, Dai *et al.* said that at shallow depth of cut, only powdery rock chips are produced, and the fluctuation amplitude of cutting force is small, while at deep cutting depth, massive rock breakage occurs, and the fluctuation amplitude of cutting force increases significantly [4]; Li *et al.* proposed that an overall positive correlation is observed between the maximum value of the cutting force and depth of cut, the cuttings size appears to increase with the increase of depth of cut [5]; Rahmani *et al.* stated that as the increase of depth of cut, the force and force area increase gradually, and the stress on the cutter changes from tensile to compressive [6]; Joodi *et al.* put forward that with the increase of depth of cut, fractures extend deeper on the rock, and the failure mode of the rock gradually changes from ductile failure to brittle failure [7]. As for the influence of wear on the performance of PDC cutter, Iman *et al.* indicated that at a wide range of depth of cut, the frictional force on the wear flat of a blunt cutter was positively related to the angle of wear flat [8]; Zhang, Yang *et al.* stated that the cutting force, temperature change and large volume cuttings of the worn cutters are more obvious compared with the new cutter, and under the same depth, the greater the wear height is, the more obvious the force fluctuation appears [9, 10]; Liu *et al.* proposed that the wear reduces the aggressiveness of cutter, and the grinding effect of cutter on cuttings is enhanced [11].

Through previous studies, the comprehensive performance and rock breaking ability of PDC cutter have been continuously improved. In particular, it is found that although the sharp cutting edge of PDC cutter can ensure its strong attack performance, when drilling into hard and heterogeneous rock, the polycrystalline diamond layer is prone to collapse. To solve this problem, some scholars put forward the scheme of pre-chamfering the edge of polycrystalline diamond layer, with the newly added chamfered surface, the stress on the cutter is extended to a larger area, thus reducing the stress gradient on the cutting edge and enhancing the impact and wear resistance of PDC cutter [12]. Shao *et al.* pointed out that the cutting efficiency can be greatly enhanced by reducing the chamfer size, but the small chamfer size makes PDC cutters vulnerable to the premature failure [13]. Fu *et al.* proposed that the chamfer has a more significant influence on the normal than the tangential cutting force, as the chamfer becomes larger, the aggressiveness of the PDC cutter decreases considerably [14]. Akbari *et al.* put forward that at shallow cuts, a chamfered PDC cutter behaves like a cutter with higher back rake angle, the deeper the cut gets the less pronounced chamfer effect [15]. Jamaludin *et al.* expressed that the significance of the cutter geometry to the wear rate is chamfer angle, back rake angle, side rake angle and cutter diameter accordingly, the chamfer angle has the strongest correlation to the wear characteristic of PDC cutter [16].

At present, the research on the influence of chamfer on the performance of PDC cutter are mostly carried out from a macro perspective, while the specific design of the chamfer parameters has not been discussed in detailed. In this work, a theoretical cutting force model of PDC cutter considering the influence of chamfer was established; based on orthogonal design test, the primary and secondary relationship of back rake angle, depth of cut, chamfer angle and chamfer length was analysed, and the influence law of chamfer parameters on cutting force were summarized. The rock cutting simulation of PDC cutters with different chamfer angles was carried out, the influence of chamfer angle on the cutting force and stress of PDC cutter were evaluated. At the same time, the drop hammer impact test of PDC cutter was carried out, so as to guide the optimization design of chamfer parameters of PDC cutter.

2. Theoretical model of cutting force considering chamfer

Decomposing the force of PDC cutter with chamfer, the resultant force F of chamfered PDC cutter is composed of the resultant force F^c on the front surface of cutter and the resultant force F^{ch} on the chamfered surface of cutter. The force decomposition diagram of PDC cutter considering the chamfer is shown in Fig. 1.

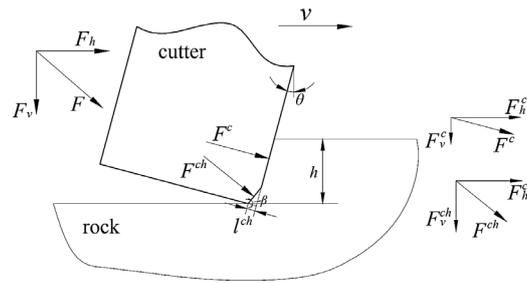


Fig. 1 Cutting force decomposition diagram of chamfered PDC cutter

According to Fig. 1, there are:

$$\begin{cases} F = \sqrt{F_h^2 + F_v^2} \\ F^c = \sqrt{F_h^{c2} + F_v^{c2}} \\ F^{ch} = \sqrt{F_h^{ch2} + F_v^{ch2}} \end{cases} \quad (1)$$

F is the resultant force of chamfered PDC cutter (N); F_h and F_v denote the horizontal force and normal force of chamfered PDC cutter (N); F^c is the resultant force on the front surface of chamfered PDC cutter (N); F_h^c and F_v^c denote the horizontal force and normal force on the front surface of chamfered PDC cutter (N); F^{ch} is the resultant force on the chamfered surface of chamfered PDC cutter (N); F_h^{ch} and F_v^{ch} denote the horizontal force and normal force on the chamfered surface of chamfered PDC cutter (N).

The force of a chamfered PDC cutter can be obtained in horizontal and vertical directions respectively:

$$\begin{cases} F_h = F_h^c + F_h^{ch} \\ F_v = F_v^c + F_v^{ch} \end{cases} \quad (2)$$

According to reference [17], it can be obtained:

$$\begin{cases} F_h^c = \varepsilon A^c \\ F_v^c = \tan(\theta + \varphi) \varepsilon A^c \end{cases} \quad (3)$$

$$\begin{cases} F_h^{ch} = \varepsilon A^{ch} \\ F_v^{ch} = \tan(\theta + \varphi) \varepsilon A^{ch} \end{cases} \quad (4)$$

A^c is the contact area between the front surface and the rock (m^2); A^{ch} is the contact area between the chamfered surface and the rock, (m^2); θ is the back rake angle (rad); ε is the intrinsic specific energy of rock (Pa); φ is the internal friction angle of rock (rad).

Substituting Eq. 3 and Eq. 4 into Eq. 1 and Eq. 2, the forces on the front surface and chamfered surface can be obtained, and the resultant force of chamfered PDC cutter can be calculated.

It is important to note that, the contact area A^c and A^{ch} are related to the relationship between chamfer angle and back rake angle, the calculation of the contact area needs to be considered from the following three cases respectively.

(1) Chamfer angle is greater than back rake angle

When the chamfer angle is greater than back rake angle, the chamfered surface of the PDC cutter is in full contact with the rock. The contact state and contact area between the PDC cutter and the rock are shown in Fig. 2.

According to Fig. 2, it can be obtained:

$$\begin{cases} A^c = S_{secOEF} - S_{triOEF} = \frac{1}{2}(\angle EOF) \cdot |OB|^2 - \frac{1}{2}|OE| \cdot |OF| \cdot \sin(\angle EOF) \\ A^{ch} = S_{secOPQ} - S_{secOEF} = \frac{1}{2}(\angle EOF) \cdot |OP|^2 - \frac{1}{2}(\angle EOF) \cdot |OB|^2 \end{cases} \quad (5)$$

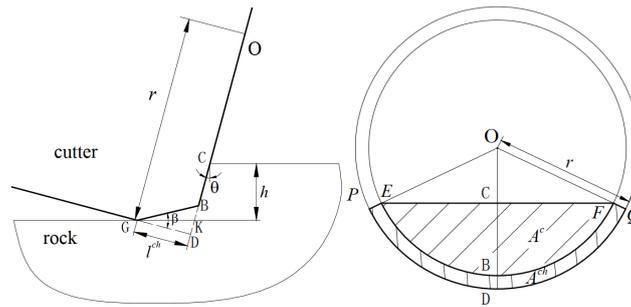


Fig. 2 Diagram of the contact state and contact area between the PDC cutter and the rock when chamfer angle is greater than back rake angle

The $\angle EOF$ is expressed in radians.

$$\begin{aligned} \angle EOF &= 2 \arccos \left(\frac{|OC|}{|OF|} \right) = 2 \arccos \left(\frac{|OD| - |BD| - |BC|}{|OB|} \right) \\ &= 2 \arccos \left[\frac{|OD| - |BD| - \left(\frac{h}{\cos \theta} - |BK| \right)}{|OD| - |BD|} \right] = 2 \arccos \left\{ \frac{|OD| - |BD| - \left[\frac{h}{\cos \theta} - (|BD| - |DK|) \right]}{|OD| - |BD|} \right\} \\ &= 2 \arccos \left\{ \frac{r - l^{ch} \tan \beta - \left[\frac{h}{\cos \theta} - (l^{ch} \tan \beta - l^{ch} \tan \theta) \right]}{r - l^{ch} \tan \beta} \right\} = 2 \arccos \left(\frac{r - \frac{h}{\cos \theta} - l^{ch} \tan \theta}{r - l^{ch} \tan \beta} \right) \end{aligned} \tag{6}$$

$$|OE| = |OF| = |OB| = |OD| - |BD| = r - l^{ch} \tan \beta \tag{7}$$

where r is the radius of chamfered PDC cutter (m); h is the depth of cut (m); β is the chamfer angle (rad); l^{ch} is the chamfer length (m).

Substituting Eq. 6 and Eq. 7 into Eq. 5, A^c and A^{ch} can be expressed:

$$\begin{cases} A^c = (r - l^{ch} \tan \beta)^2 \cdot \left\{ \arccos \left(\frac{r - \frac{h}{\cos \theta} - l^{ch} \tan \theta}{r - l^{ch} \tan \beta} \right) - \frac{1}{2} \sin \left[2 \arccos \left(\frac{r - \frac{h}{\cos \theta} - l^{ch} \tan \theta}{r - l^{ch} \tan \beta} \right) \right] \right\} \\ A^{ch} = \arccos \left(\frac{r - \frac{h}{\cos \theta} - l^{ch} \tan \theta}{r - l^{ch} \tan \beta} \right) \cdot [r^2 - (r - l^{ch} \tan \beta)^2] \end{cases} \tag{8}$$

Substituting Eq. 8 into Eq. 3 and Eq. 4, the horizontal force and normal force acting on the front surface and chamfered surface of PDC cutter can be obtained.

(2) Chamfer angle is equal to back rake angle

Fig. 3 shows the contact state and contact area between the PDC cutter and the rock when the chamfer angle is equal to back rake angle, at this time, the chamfered line BG showed in Fig. 3 is parallel to the horizontal direction.

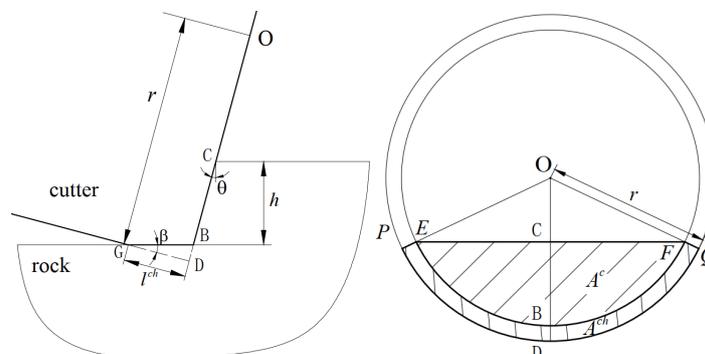


Fig. 3 Diagram of the contact state and contact area between the PDC cutter and rock when chamfer angle is equal to back rake angle

According to Fig. 3, it can be obtained:

$$\begin{cases} F_h^{ch} = \mu F_v^{ch} \\ F_v^{ch} = \tan(\theta + \varphi) \varepsilon A^{ch} \end{cases} \quad (9)$$

where μ is the friction coefficient of rock.

$$\begin{cases} A^c = S_{secOEF} - S_{triOEF} = \frac{1}{2}(\angle EOF) \cdot |OB|^2 - \frac{1}{2}|OE| \cdot |OF| \cdot \sin(\angle EOF) \\ A^{ch} = S_{secOPQ} - S_{secOEF} = \frac{1}{2}(\angle EOF) \cdot |OP|^2 - \frac{1}{2}(\angle EOF) \cdot |OB|^2 \end{cases} \quad (10)$$

$$\begin{aligned} \angle EOF &= 2 \arccos\left(\frac{|OC|}{|OF|}\right) = 2 \arccos\left(\frac{|OD| - |BD| - |BC|}{|OB|}\right) \\ &= 2 \arccos\left(\frac{|OD| - |BD| - |BC|}{|OD| - |BD|}\right) = 2 \arccos\left(\frac{r - l^{ch} \tan \beta - \frac{h}{\cos \theta}}{r - l^{ch} \tan \beta}\right) \end{aligned} \quad (11)$$

$$|OE| = |OF| = |OB| = |OD| - |BD| = r - l^{ch} \tan \beta \quad (12)$$

Substituting Eq. 11 and Eq. 12 into Eq. 10, A^c and A^{ch} can be expressed:

$$\begin{cases} A^c = (r - l^{ch} \tan \beta)^2 \cdot \left\{ \arccos\left(\frac{r - l^{ch} \tan \beta - \frac{h}{\cos \theta}}{r - l^{ch} \tan \beta}\right) - \frac{1}{2} \sin\left[2 \arccos\left(\frac{r - l^{ch} \tan \beta - \frac{h}{\cos \theta}}{r - l^{ch} \tan \beta}\right)\right] \right\} \\ A^{ch} = \arccos\left(\frac{r - l^{ch} \tan \beta - \frac{h}{\cos \theta}}{r - l^{ch} \tan \beta}\right) \cdot [r^2 - (r - l^{ch} \tan \beta)^2] \end{cases} \quad (13)$$

Substituting Eq. 13 into Eq. 3 and Eq. 9, the horizontal force and normal force acting on the front surface and chamfered surface of PDC cutter can be obtained.

(3) Chamfer angle is less than back rake angle

When the chamfer angle is less than back rake angle, in an ideal state, the chamfered surface of the PDC cutter does not contact with the rock. The contact state and contact area between the PDC cutter and the rock are shown in Fig. 4.

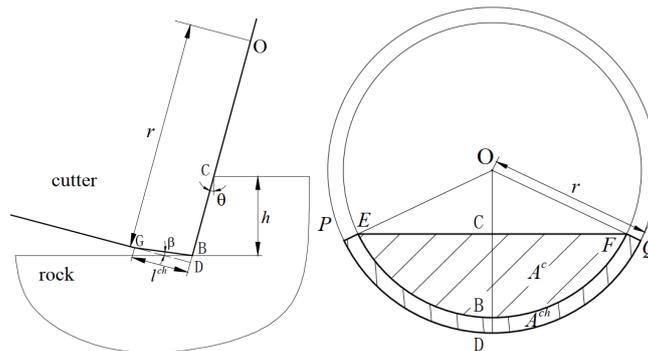


Fig. 4 Diagram of the contact state and contact area between the PDC cutter and rock when chamfer angle is equal to back rake angle

According to Fig. 4, it can be obtained:

$$F_h^{ch} = F_v^{ch} = 0 \quad (14)$$

$$A^c = S_{secOEF} - S_{triOEF} = \frac{1}{2}(\angle EOF) \cdot |OB|^2 - \frac{1}{2}|OE| \cdot |OF| \cdot \sin(\angle EOF) \quad (15)$$

$$\begin{aligned} \angle EOF &= 2 \arccos\left(\frac{|OC|}{|OF|}\right) = 2 \arccos\left(\frac{|OD| - |BD| - |BC|}{|OB|}\right) \\ &= 2 \arccos\left(\frac{|OD| - |BD| - |BC|}{|OD| - |BD|}\right) = 2 \arccos\left(\frac{r - l^{ch} \tan \beta - \frac{h}{\cos \theta}}{r - l^{ch} \tan \beta}\right) \end{aligned} \quad (16)$$

$$|OE| = |OF| = |OB| = |OD| - |BD| = r - l^{ch}\tan \beta \tag{17}$$

Substituting Eq. 16 and Eq. 17 into Eq. 15, A^c can be expressed:

$$A^c = (r - l^{ch}\tan \beta)^2 \cdot \left\{ \arccos \left(\frac{r - l^{ch}\tan \beta - \frac{h}{\cos \theta}}{r - l^{ch}\tan \beta} \right) - \frac{1}{2} \sin \left[2 \arccos \left(\frac{r - l^{ch}\tan \beta - \frac{h}{\cos \theta}}{r - l^{ch}\tan \beta} \right) \right] \right\} \tag{18}$$

Substituting Eq. 18 into Eq. 3, the horizontal force and normal force acting on the front surface of PDC cutter can be obtained.

3. Primary and secondary factors affecting the force of cutter

Based on previous research, it is known that back rake angle and depth of cut have significant influence on the force of cutter [18-23]. Therefore, combined with the focus of this work, the orthogonal design test of four factors, i.e. back rake angle, depth of cut, chamfer angle and chamfer length, was established, the influence degree of the four factors on the force of cutter was analysed according to the above established equations.

3.1 Pseudo-level orthogonal test design

Table 1 is the orthogonal design level table of the factors affecting the force of cutter.

According to Table 1, orthogonal test table $L_{81}(4^9)$ was selected for the pseudo-level design of the four factors. According to experience, the designed back rake angle is generally between 10 and 20° [24, 25], the depth of cut that the bit drill into the rock per turn is generally between 2 and 3 mm [26, 27], therefore, in the pseudo-level orthogonal test table, the second, third and fourth levels (10°, 15° and 18°) of the back rake angle were repeated once respectively, and the third, fourth and fifth levels (2 mm, 2.5 mm, 3 mm) of the depth of cut were repeated once respectively. Because the influence of chamfer length on the force of cutter is unknown, a random repetition of the chamfer length level was performed. Based on this, the pseudo-level orthogonal test table was designed, and the parameters in the table were substituted into the equations derived in Section 2, the results are shown in Table 2.

Table 1 Orthogonal design level table of the factors affecting the force of cutter

Factor Level	A Back rake angle (°)	B Cutting depth (°)	C Chamfer angle (°)	D Chamfer length (°)
1	5	1	5	0.2
2	10	1.5	10	0.3
3	15	2	15	0.4
4	18	2.5	18	0.5
5	20	3	20	
6	25	3.5	25	
7			30	
8			45	
9			60	

Table 2 Pseudo-level orthogonal test level table of the force influencing factors and the results

No.	Back rake angle (°)	Depth of cut (mm)	Chamfer angle (°)	Chamfer length (mm)	Resultant force on front surface (N)	Resultant force on chamfered surface (N)	Resultant force of PDC cutter (N)
1	25	1.5	60	0.3	878.381	241.014	1119.395
2	5	2.5	45	0.4	1695.374	256.402	1951.776
3	10	3	45	0.2	1603.419	130.957	1734.376
4	18	1.5	25	0.4	737.354	130.366	867.719
5	15	3.5	18	0.2	2378.183	70.499	2448.682
6	18	2	60	0.4	978.008	297.613	1275.620
7	5	2	15	0.3	828.478	53.342	881.820
8	15	2	45	0.3	963.713	172.976	1136.689
9	18	2	18	0.3	1169.110	80.727	1249.067
10	15	1	10	0.5	388.434	0.000	388.434
11	18	2.5	15	0.2	1618.022	0.000	1618.022
12	15	2.5	20	0.3	1457.285	97.073	1554.358
13	15	1.5	18	0.2	700.775	44.704	745.479

Table 2 (continuation)

No.	Back rake angle (°)	Depth of cut (mm)	Chamfer angle (°)	Chamfer length (mm)	Resultant force on front surface (N)	Resultant force on chamfered surface (N)	Resultant force of PDC cutter (N)
14	20	3	5	0.2	2245.877	0.000	2245.877
15	10	2	60	0.2	848.894	126.749	975.643
16	18	2.5	5	0.5	1619.379	0.000	1619.379
17	18	1	20	0.3	420.950	64.812	485.761
18	18	3	5	0.3	2102.773	0.000	2102.773
19	25	3.5	5	0.3	3363.261	0.000	3363.261
20	18	3	10	0.5	2091.386	0.000	2091.386
21	18	3	25	0.2	2073.736	95.201	2168.936
22	18	3	60	0.4	1847.717	377.403	2225.120
23	10	2	25	0.5	855.799	155.301	1011.099
24	5	3	60	0.5	1225.042	348.676	1573.718
25	10	1.5	5	0.3	625.856	0.000	625.856
26	15	2.5	5	0.5	1482.830	0.000	1482.830
27	25	2.5	10	0.4	2083.419	0.000	2083.419
28	20	3.5	10	0.3	2771.957	0.000	2771.957
29	25	2	20	0.4	1502.580	0.000	1502.580
30	10	3.5	20	0.5	2021.925	173.535	2195.459
31	18	2.5	45	0.5	1414.735	348.157	1762.892
32	18	3.5	15	0.4	2593.402	0.000	2593.402
33	25	1	45	0.2	511.444	108.000	619.444
34	18	2.5	20	0.2	1612.969	69.961	1682.929
35	10	1	15	0.4	326.428	53.884	380.312
36	15	3	10	0.4	1918.500	0.000	1918.500
37	15	3	30	0.3	1844.863	155.515	2000.378
38	15	3.5	45	0.4	2180.744	314.904	2495.649
39	10	3	18	0.5	1632.464	144.143	1776.607
40	20	2.5	15	0.2	1727.147	0.000	1727.147
41	25	2.5	25	0.5	2054.058	291.706	2345.709
42	25	3	18	0.5	2676.756	0.000	2676.756
43	20	2	45	0.5	1081.393	324.940	1406.333
44	10	2.5	10	0.3	1310.027	36.491	1345.110
45	10	3.5	60	0.2	1976.607	174.213	2150.819
46	15	2.5	60	0.3	1312.954	237.843	1550.797
47	15	3	30	0.4	1816.806	206.570	2023.375
48	10	3	45	0.3	1552.690	194.604	1747.294
49	10	1	18	0.4	316.120	63.660	379.781
50	10	3	15	0.3	1676.193	72.944	1749.138
51	20	3	20	0.3	2227.834	121.871	2349.031
52	5	3	20	0.4	1458.881	116.273	1575.154
53	18	1	30	0.3	391.643	92.531	484.174
54	20	2.5	18	0.4	1715.850	0.000	1715.850
55	10	2.5	30	0.2	1257.302	84.328	1341.630
56	18	2	10	0.2	1175.170	0.000	1175.170
57	5	2.5	18	0.3	1140.615	71.677	1212.292
58	10	2	10	0.3	949.537	32.402	980.695
59	15	2	5	0.4	1075.850	0.000	1075.850
60	10	2.5	25	0.4	1225.452	141.568	1367.020
61	20	2	25	0.3	1230.262	121.130	1351.392
62	5	1	5	0.2	313.160	6.234	318.981
63	5	1.5	10	0.2	558.197	20.674	578.871
64	15	3	25	0.2	1889.151	88.195	1977.347
65	5	2	30	0.5	724.446	164.589	889.035
66	18	3.5	30	0.5	2474.827	304.077	2778.904
67	20	1	60	0.5	296.278	252.482	548.761
68	5	3.5	25	0.3	1829.778	117.601	1947.379
69	15	1	25	0.3	361.907	72.841	434.748
70	25	3	15	0.3	2696.483	0.000	2696.483
71	10	2	5	0.4	950.940	0.000	950.940
72	15	2	15	0.5	1067.024	98.846	1164.149
73	20	1.5	30	0.4	774.430	162.056	936.486
74	10	1.5	20	0.5	569.941	108.467	678.408
75	15	2	20	0.2	1061.781	57.489	1119.269
76	15	1.5	15	0.5	702.100	84.981	785.639
77	25	2	30	0.2	1501.723	112.019	1613.742
78	10	2.5	30	0.3	1228.481	125.641	1354.122
79	18	1.5	45	0.3	682.543	159.383	841.926

3.2 Multi-factor variance analysis

In the drilling process, rock breaking force is simultaneously affected by four factors: back rake angle, depth of cut, chamfer angle and chamfer length. Therefore, taking the resultant force on the front surface, the resultant force on the chamfered surface, and the resultant force of PDC cutter as targets respectively, the data in Table 2 were analysed by multi-factor variance method [28]. Taking the resultant force on the front surface of PDC cutter as object, the varication analysis results of each factor are shown in Table 3.

Table 3 Multi-factor variance analysis with the object of the resultant force on the front surface of PDC cutter

Source of variation	Sum of deviation square	Freedom	Mean square	Value of F
Factor A (back rake angle)	12303791.19	5	2460758.239	2.145812091
Factor B (depth of cut)	53507230.5	5	10701446.1	9.331795407
Factor C (chamfer angle)	1508953.542	8	188619.1928	0.164478305
Factor D (chamfer length)	3924539.023	3	1308179.674	1.140749107
Factor E (error)	9174179.787	8	1146772.473	

According to Table 3, from high to low, the influence degree of each factor on the resultant force of PDC cutter is as follows: depth of cut, back rake angle, chamfer length and chamfer angle. The depth of cut shows the most significant effect on the resultant force of front surface.

Taking the resultant force on the chamfered surface of PDC cutter as object, the varication analysis results of each factor are shown in Table 4.

Table 4 Multi-factor variance analysis with the object of the resultant force on the chamfered surface of PDC cutter

Source of variation	Sum of deviation square	Freedom	Mean square	Value of F
Factor A (back rake angle)	99076.91606	5	19815.38321	1.331109649
Factor B (depth of cut)	109763.2665	5	21952.6533	1.474681984
Factor C (chamfer angle)	551868.6539	8	68983.58174	4.634011378
Factor D (chamfer length)	69317.13208	3	23105.71069	1.552139271
Factor E (error)	119090.9147	8	14886.36434	

According to Table 4, from high to low, the influence degree of each factor on the resultant force of PDC cutter is as follows: chamfer angle, chamfer length, depth of cut and back rake angle. The chamfer angle shows the most significant effect on the resultant force of chamfered surface.

Taking the resultant force of PDC cutter as object, the varication analysis results of each factor are shown in Table 5.

Table 5 Multi-factor variance analysis with the object of the resultant force of PDC cutter

Source of variation	Sum of deviation square	Freedom	Mean square	Value of F
Factor A (back rake angle)	14082530.17	5	2816506.033	2.320851882
Factor B (depth of cut)	57888672.68	5	11577734.54	9.540262536
Factor C (chamfer angle)	568633.273	8	71079.15912	0.058570512
Factor D (chamfer length)	4160716.972	3	1386905.657	1.142835331
Factor E (error)	9708524.89	8	1213565.611	

According to Table 5, from high to low, the influence degree of each factor on the resultant force of PDC cutter is as follows: depth of cut, back rake angle, chamfer length and chamfer angle. The depth of cut shows the most significant effect on the resultant force of PDC cutter, while the chamfer angle and chamfer length have no significant effect.

Through multi-factor variance analysis, it can be obtained that, the depth of cut plays a major role in the force of PDC cutter, followed by the back rake angle. The chamfer shows the significant effect on the resultant force of chamfered surface, it mainly affects the stress distribution of the PDC cutter, that is, the chamfer mainly affects the wear and impact resistance of the cutter.

4. Influence of chamfer parameters on the force of PDC cutter

Although the chamfer contributes minimally to the force of the PDC cutter, it is essential to examine the influence of the chamfer parameters on the force. This examination will provide guidance for the design of chamfer parameters. Based on the equations established in Section 2, the influence of the chamfer angle and chamfer length on the force of the PDC cutter was analysed separately.

4.1 Influence of chamfer angle on the force of PDC cutter

Based on previous research, fixing the back rake angle 15° and the depth of cut 0.3 mm unchanged [29, 30], the influence law of chamfer angle on the force of PDC cutter with different chamfer lengths was analysed. Figs. 5, 6 and 7 show the variation characteristics of the resultant force on front surface, the resultant force on chamfered surface and the resultant force of PDC cutter with chamfer angle respectively.

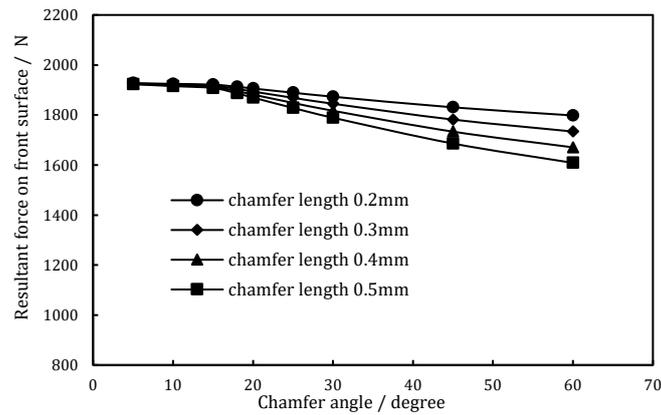


Fig. 5 Variation curve of the resultant force on front surface of PDC cutter with chamfer angle

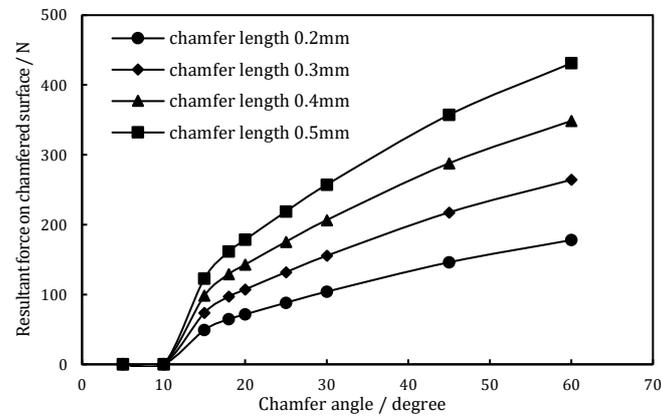


Fig. 6 Variation curve of the resultant force on chamfered surface of PDC cutter with chamfer angle

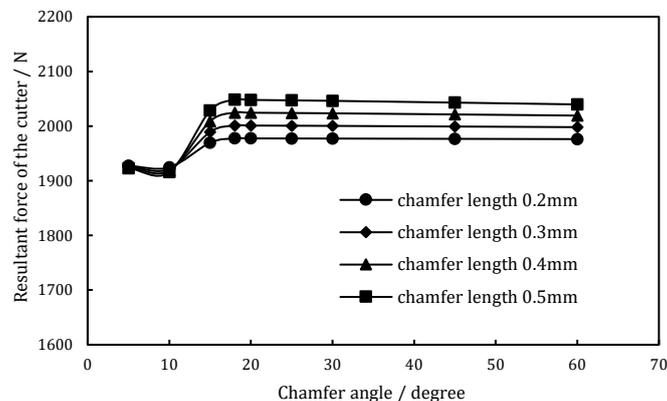


Fig. 7 Variation curve of the resultant force of PDC cutter with chamfer angle

According to Figs. 5, 6 and 7, when the chamfer angle is small (chamfer angle not greater than back rake angle), the chamfer angle appears almost no influence on the resultant force on both front and chamfered surface, while the resultant force of PDC cutter increases significantly with the increase of chamfer angle. When the chamfer angle is greater than back rake angle, with the increase of chamfer angle, the resultant force on the front surface decreases linearly, while the resultant force on chamfered surface increases approximately linearly, and the larger the chamfer length, the faster the reduction or increase rate. The resultant force of PDC cutter increases first and then decreases slightly with the increase of chamfer angle, the decreasing trend can be ignored when the chamfer length is small.

4.2 Influence of chamfer length on the force of PDC cutter

Fixing the back rake angle 15° and the depth of cut 0.3 mm unchanged, the influence of chamfer length on the force of PDC cutter with different chamfer angles was analysed. Figs. 8, 9 and 10 show the variation characteristics of the resultant force on front surface, the resultant force on chamfered surface and the resultant force of PDC cutter with chamfer length respectively.

According to Figs. 8, 9 and 10, when the chamfer angle is small (chamfer angle not greater than back rake angle), the chamfer length appears almost no influence on both the resultant force on front surface, chamfered surface and the PDC cutter. When chamfer angle is greater than back rake angle, with the increase of chamfer length, the resultant force on the front surface decreases linearly, while the resultant force on chamfered surface increases approximately linearly, and the greater the chamfer angle, the faster the reduction or increase rate. The resultant force of PDC cutter increases linearly with the increase of chamfer length, and the amplitude of linear increase is almost unchanged under different chamfer angles.

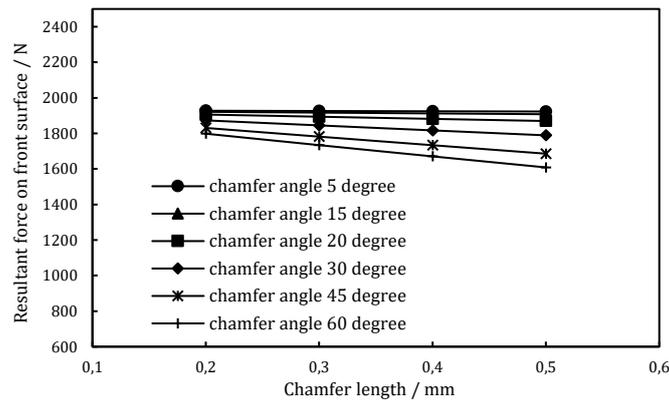


Fig. 8 Variation curve of the resultant force on front surface of PDC cutter with chamfer length

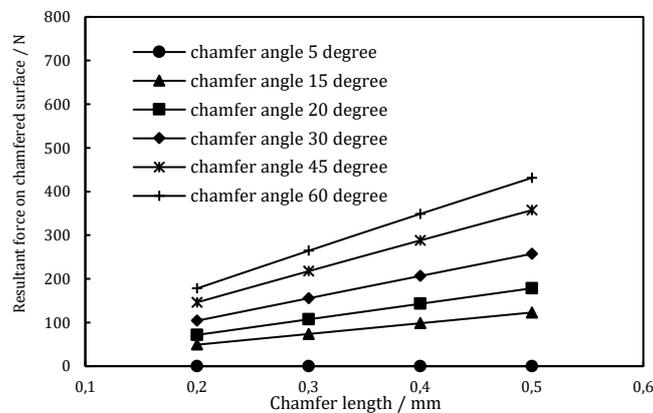


Fig. 9 Variation curve of the resultant force on chamfered surface of PDC cutter with chamfer length

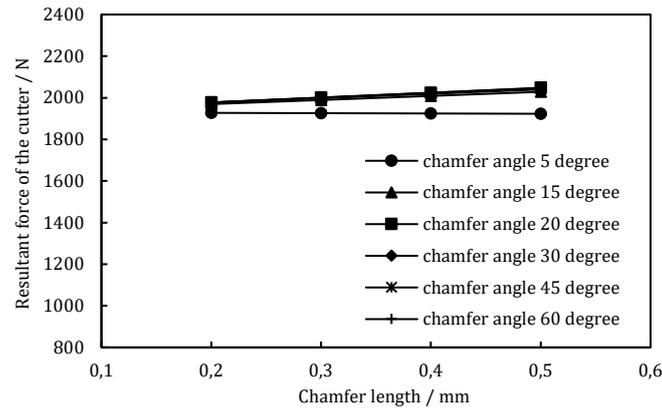


Fig. 10 Variation curve of the resultant force of PDC cutter with chamfer length

5. Rock breaking simulation of PDC cutters with different chamfer angles

It is beneficial to improve the stress distribution by chamfering the PDC cutter, but it is difficult to observe the stress state of the PDC cutter through theoretical and experimental analysis. So in this work, numerical simulation based on Smooth Particle Hydrodynamic (SPH) method was used to reveal the rock breaking characteristics and stress state of PDC cutters with different chamfer angles.

5.1 Simulation model

The SPH is a meshless continuum mechanics method, in which the computation domain is discretized into a series of interacting particles. Its good adaptive characteristics make it can deal with the problems of large deformation and post-instability well, so can effectively avoid the mesh distortion caused by rock breaking [31, 32]. In this work, the following assumptions were made: (1) the rocks were continuous, homogeneous and isotropic; (2) the influence of temperature and fluid was ignored; (3) regardless of repeated breaking, the rock unit was deleted immediately after being broken.

Drucker-Prager model was used to simulate the constitutive relation of rock [33]. In this model, the formation and peeling process of cuttings can be represented by setting shear failure criterion, and the shear failure of rock can be simulated well. The failure equivalent plastic strain was defined as the criterion of rock damage. When the equivalent plastic strain of rock was equal to the failure plastic strain, the rock element was broken and deleted [24]. The cutter was represented by an elastic model, which can consider the wear of the cutter. Table 6 shows the material parameters of the cutter and the rock.

The size of rock was $40 \times 30 \times 30$ mm, and it was modelled by SPH particles, the particle size was 0.8 mm. The size of cutter was $\Phi 13.44 \times 8$ mm, and it was modelled by hexahedral element, the mesh size was 0.4 mm. The rock was completely fixed, the cutter cut the rock linearly along the Y axis at a speed of 100 m/h, and the depth of cut was set 3 mm unchanged. Fig. 11 shows the 3D model of linear cutting rock by PDC cutter.

The idea of unit erosion algorithm was taken for reference to deal with the failure of SPH rock particles, the rock particles were deemed to be invalid when the damage variable of rock particles reach the critical value, the rock particles were deleted, but the mass and momentum of the failed particles were still retained, thus ensuring the conservation of mass and momentum of the system.

Table 6 Material parameters of the cutter and rock

	Density ($\text{kg}\cdot\text{m}^{-3}$)	Elastic modulus (GPa)	Poisson's ratio	Compressive strength (MPa)	Shear strength (MPa)	Cohesion (MPa)	Internal friction angle ($^{\circ}$)
Cutter	3560	850	0.07				
Rock	2500	40	0.27	75	10.0	27.2	35.0

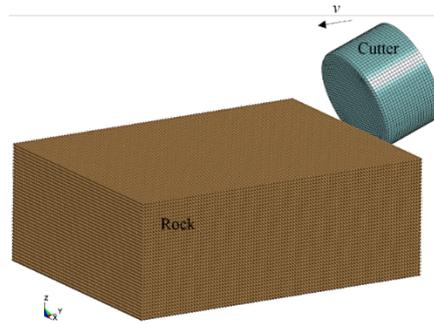


Fig. 11 3D model of linear cutting rock by PDC cutter

According to the analysis in Section 2, the relationship between the chamfer angle and back rake angle determines the contact state between PDC cutter and the rock. Therefore, in this simulation, the rock breaking characteristics of PDC cutters with different chamfer angles under different back rake angles were analysed. Table 7 shows the parameters of PDC cutter in the simulation.

Table 7 Parameters of PDC cutter

Cutter size (mm)	Chamfer length (mm)	Back rake angle (°)	Chamfer angle (°)
Φ 13.44 × 8	0.3	15	15
			30
			45
			60
		18	15
			30
			45
			60
		20	15
			30
			45
			60

5.2 Rock breaking force

Fig. 12 shows the stress nephogram of rock and force curve of PDC cutter during rock breaking process. As can be observed, the rock particles were broken under the shear and extrusion actions of PDC cutter, an arc-shaped crushing pit was formed, the maximum stress on the rock was located in the front surface of the cutting edge, which took the shape of an arc belt. When the rock breaking process was stable, the cutting force, axial force and lateral force all fluctuated with time, and the collapse failure of rock particles was reflected in the sudden decrease of cutting force. In this work, all the analysis are based on the stable cutting stage, and the lateral force is relatively small, so it is ignored.

Fig. 13 shows the force curves of PDC cutters with different chamfer angles under different back rake angles. As can be observed, both the cutting force and the axial force increase with the increase of back rake angle, which is consistent with the conclusion in reference 34 [34]. Under all the back rake angles, the cutting force and axial force present the same changing law with the increase of chamfer angle, that is, a trend of first increasing and then slightly decreasing, and the decrease amplitude is very small (which is consistent with the conclusion obtained in Fig. 7). Moreover, the larger the back rake angle is, the more significant the change is. When the back rake angle is 15°, the cutting force and axial force show little difference with the chamfer angle. When the chamfer angle is less than back rake angle, the influence of chamfer angle on cutting force and axial force is more obvious, while when the chamfer angle is larger than back rake angle, the influence of chamfer angle on cutting force and axial force is very small, which can be almost ignored.

According to Fig. 12, the cutting force fluctuates up and down around a certain value in the process of rock breaking, the more severe the fluctuation, the greater the possibility of vibration damage to the PDC cutter. The cutting force difference coefficient is used to evaluate the severity of the fluctuation, which is defined as the ratio between the standard deviation of cutting force

and the mean value of cutting force. The smaller the cutting force difference coefficient is, the gentler the fluctuation of the cutting force is. Fig. 14 shows the cutting force difference coefficient curves of PDC cutters with different chamfer angles under different back rake angles.

It can be seen from Fig. 14, when the chamfer angle is between 30° and 45°, the cutting force difference coefficient is small, and the rock breaking process is relatively stable. Similar conclusions can be drawn at different back rake angles.

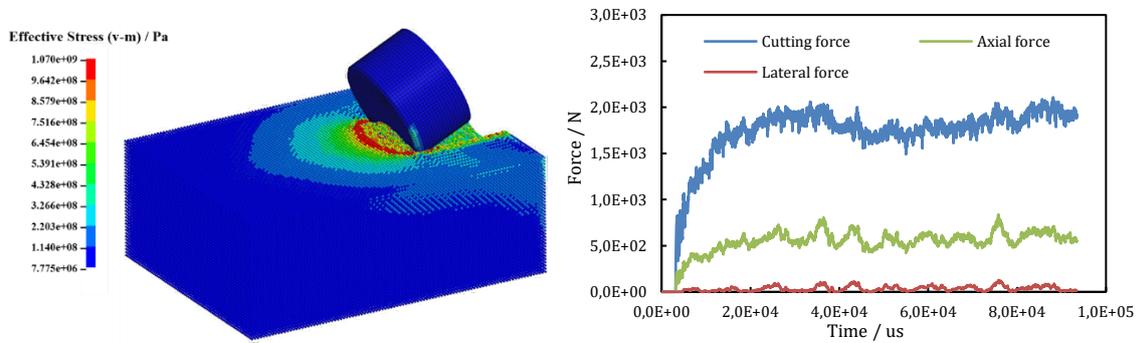


Fig. 12 Stress nephogram of rock and force curve of PDC cutter

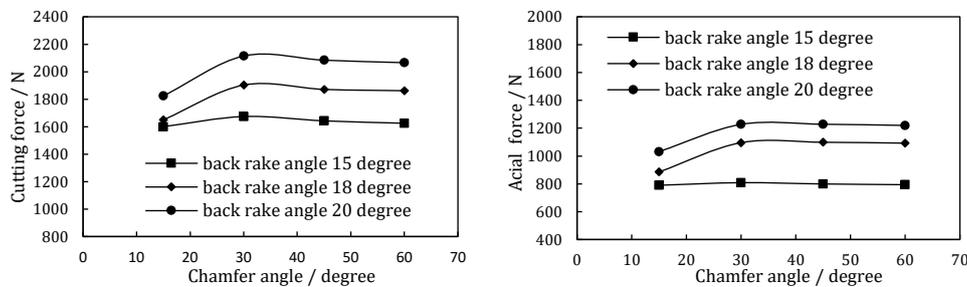


Fig. 13 Force curves of PDC cutters with different chamfer angles under different back rake angles

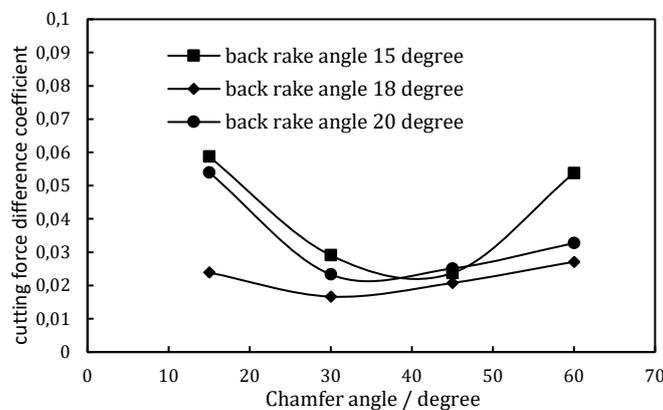


Fig. 14 Cutting force difference coefficient curves of PDC cutters with different chamfer angles

5.3 Stress distribution characteristic of PDC cutter

Fig. 15 shows the stress distribution nephogram of PDC cutters with different chamfer angles at a back rake angle of 25°. With the increase of chamfer angle, the maximum stress on the PDC cutter decreases gradually, and the stress distribution area expands from the cutting edge to a larger area. Specifically, when the chamfer angle is 15°, the surface stress on the PDC cutter is concentrated in the part where the lower edge of the PDC cutter contacts with the rock. At this time, the stress on the lower of the cutting edge is concentrated, and the collapse failure of the cutting edge is easy to occur. When the chamfer angle is increased to 60°, the surface stress on the PDC cutter spreads to a larger area, which effectively protects the cutting edge, prevents the occurrence of early cracks and spalls of the polycrystalline diamond layer.

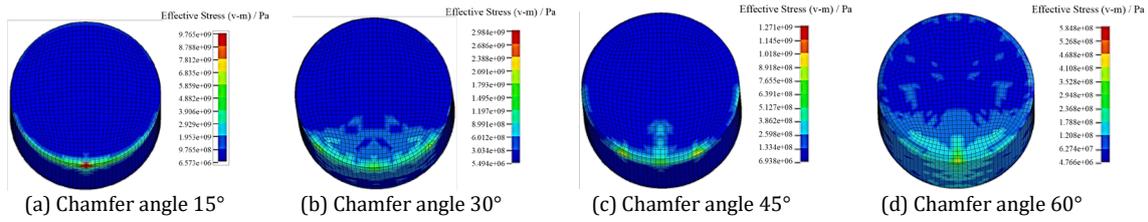


Fig. 15 Stress distribution nephogram of PDC cutters with different chamfer angles

6. Impact resistance test of PDC cutters with different chamfer angles

Drop hammer impact tests of PDC cutters with chamfer angles of 15°, 30°, 45° and 60° were carried out to further verify the influence of chamfer on the impact resistance of PDC cutter. Fig. 16 shows the photos of PDC cutters with different chamfer angles, and the cutter sizes are $\Phi 13.44 \times 8$ mm.

The drop hammer impact device was used in the test (shown as Fig. 17), PDC cutters were fixed on the punch head with a back rake angle of 20°. The mould steel was continuously impacted by the PDC cutter with a given energy, the hardness of the mould steel was HRC58-62. Every time the punch impacted, the mould steel rotated by 10°, so to ensure the PDC cutter can impact the undamaged mould steel plane every time.

The impact energy was loaded in a step-by-step energy mode. Starting from 10 J impact energy for 10 times, the energy was increased step-by-step for 20 J impact energy for 10 times, 30 J impact energy for 10 times and 40 J impact energy for 10 times, the test was stopped once the polycrystalline diamond layer was damaged. The total impact energy that the PDC cutter can withstand was calculated according to the sum of the single impact energy. Two pieces of each chamfered cutter were tested, the average of the two test results was taken as the final impact resistance energy of the chamfered cutter. Table 8 shows the cumulative average impact energy of the PDC cutter with different chamfer angles.



Fig. 16 Photos of PDC cutters with different chamfer angles

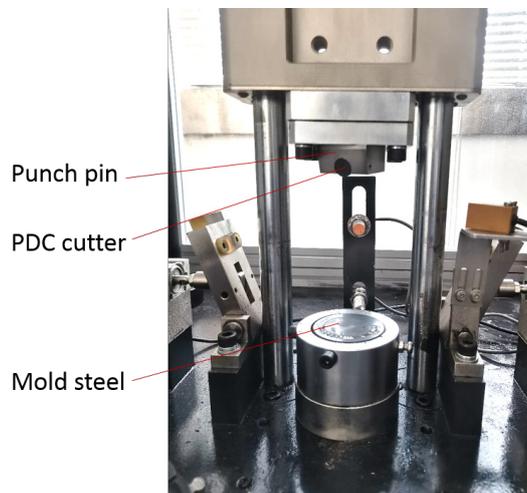


Fig. 17 Photo of drop hammer impact device

Table 8 Impact test data of PDC cutters

PDC cutter	Number of impact				Cumulative average impact energy/ J	Damage morphology
	10J	20J	30J	40J		
Chamfer angle 15°	10	10	2	/	500	
	10	10	10	1		
Chamfer angle 30°	10	10	10	1	605	
	10	10	9	/		
Chamfer angle 45°	10	10	10	2	740	
	10	10	10	5		
Chamfer angle 60°	10	10	10	10	920	
	10	10	10	6		

According to Table 8, when the impact energy was less than 20 J, all the PDC cutters were intact. When the impact energy was increased to 30 J, the whole polycrystalline diamond layer collapsed on one PDC cutter with chamfer angle of 15°, and the side edge of polycrystalline diamond layer broke on one PDC cutter with chamfer angle of 30°. When the impact energy was increased to 40 J, the main damage forms of all the PDC cutters were the whole collapsing of polycrystalline diamond layer, the upper surface of polycrystalline diamond layer cracked on one PDC cutter with chamfer angle of 45°, and the side edge of polycrystalline diamond layer broke on one PDC cutter with chamfer angle of 60°. Considering the impact energy and damage morphology comprehensively, the impact resistance of PDC cutters with chamfer angle of 45° and 60° were superior, and this was in good agreement with the simulation results.

7. Conclusion

In this study, theoretical model of cutting force considering chamfer was established, the primary and secondary relationship of four factors – back rake angle, depth of cut, chamfer angle and chamfer length – on the force of PDC cutter was analysed through pseudo-level orthogonal level test. The depth of cut played a major role in the resultant force on front surface and the PDC cutter, followed by the back rake angle. The chamfer angle showed the most significant effect on the resultant force on chamfered surface.

The influence of the chamfer parameters on the force of PDC cutter was analysed. When the chamfer angle was small (chamfer angle not greater than back rake angle), the chamfer angle and chamfer length appeared almost no influence on the resultant force on both front surface and chamfered surface of PDC cutter, while the resultant force of PDC cutter increased significantly with the increase of chamfer angle. When chamfer angle was greater than back rake angle, with the increase of chamfer angle and chamfer length, the resultant force on the front surface decreased linearly, while the resultant force on the chamfered surface increased approximately linearly; for the resultant force of PDC cutter, it increased first and then decreased slightly with the increase of chamfer angle, also it increased linearly with the increase of chamfer length.

Numerical simulation based on SPH method was carried out to analyse the rock breaking characteristics of PDC cutters with different chamfer angles. With the increase of chamfer angle, both cutting force and axial force increased first and then decreased slightly, also the decrease amplitude was very small. When the chamfer angle was less than back rake angle, the influence of chamfer angle on cutting force and axial force was more obvious. When the chamfer angle was between 30° and 45°, the rock breaking process was relatively stable. With the increase of chamfer angle,

the maximum stress on the PDC cutter decreased gradually, and the distribution area of stress on the PDC cutter was wider, which was beneficial to protect the cutting edge, and prevent the occurrence of early cracks and spalls in the polycrystalline diamond layer of PDC cutter. The same results were obtained through drop hammer impact tests.

This work can provide a theoretical basis for the optimal design of chamfer parameters, future research will focus on studying the rock breaking effect of chamfered PDC cutters with different sizes (Φ 16 mm, Φ 19 mm), aiming to offer more comprehensive and perfect guidance for the optimization design of chamfer parameters. At the same time, field drilling test of PDC bit with different chamfered PDC cutters can be carried out, combined with the analysis of micro-morphology, the micro-mechanism of the chamfer on improving the impact resistance of the PDC cutter can be further revealed.

Acknowledgements

This research was supported by Key Research and Development project in Shaanxi Province, grant number 2022GY-146, Science and Technology Innovation Fund Project of CCTEG Xi'an Research Institute, grant number 2021XAYJCQ02, and the National Key Research and Development project for young scientists, grant number 2021YFC2900200. I appreciate their support very much.

References

- [1] Rajabov, V., Miska, S., Mortimer, L., Yu, M., Ozbayoglu, E. (2012). The effects of back rake and side rake angles on mechanical specific energy of single PDC cutters with selected rocks at varying depth of cuts and confining pressures, In: *Proceedings of the IADC/SPE Drilling Conference and Exhibition*, San Diego, California, USA, 1-17, [doi: 10.2118/151406-MS](https://doi.org/10.2118/151406-MS).
- [2] Zhang, Z., Zhao, D., Zhao, Y., Gao, K., Zhang, C., Lü, X. (2023). 3D numerical simulation study of rock breaking of the wavy PDC cutter and field verification, *Journal of Petroleum Science and Engineering*, Vol. 203, Article No. 108578, [doi: 10.1016/j.petrol.2021.108578](https://doi.org/10.1016/j.petrol.2021.108578).
- [3] Adzis, A.H.A., Abdul-Rani, A.M., Yi, K.Y., Maulianda, B.T., Rao, T.V.V.L.N. (2018). Effect of back rake angle and shape on wear rate of PDC cutter in hard formation, In: *Proceedings of the 6th International Conference on Production, Energy and Reliability, World Engineering Science & Technology Congress (ESTCON)*, Kuala Lumpur, Malaysia, 1-9, [doi: 10.1063/1.5075585](https://doi.org/10.1063/1.5075585).
- [4] Dai, X.-W., Huang, Z.-W., Huang, P.-J., Chen, P.-J., Shi, H.-Z., Yan, S. (2023). Experimental investigation on the cuttings formation process and its relationship with cutting force in single PDC cutter tests, *Petroleum Science*, Vol. 20, No. 3, 1779-1787, [doi: 10.1016/j.petsci.2022.10.021](https://doi.org/10.1016/j.petsci.2022.10.021).
- [5] Li, Y., Chen, Z., Ye, Y., Yang, Y. (2021). Combined finite-discrete element method for modeling the interaction between single PDC cutter and brittle rock, *Journal of Petroleum Science and Engineering*, Vol. 207, Article No. 109133, [doi: 10.1016/j.petrol.2021.109133](https://doi.org/10.1016/j.petrol.2021.109133).
- [6] Rahmani, R., Pastusek, P., Yun, G., Roberts, T. (2020). Investigation of geometry and loading effects on PDC cutter structural integrity in hard rocks, In: *Proceedings of the IADC/SPE International Drilling Conference and Exhibition*, Galveston, Texas, USA, 1-22, [doi: 10.2118/199598-MS](https://doi.org/10.2118/199598-MS).
- [7] Joodi, B., Sarmadivaleh, M., Rasouli, V., Nabipour, A. (2012). Simulation of the cutting action of a single PDC cutter using DEM, *Petroleum and Mineral Resources*, Vol. 81, 143-150, [doi: 10.2495/PMR120131](https://doi.org/10.2495/PMR120131).
- [8] Rostamsowlat, I., Akbari, B., Evans, B. (2018). Analysis of rock cutting process with a blunt PDC cutter under different wear flat inclination angles, *Journal of Petroleum Science and Engineering*, Vol. 171, 771-783, [doi: 10.1016/j.petrol.2018.06.003](https://doi.org/10.1016/j.petrol.2018.06.003).
- [9] Zhang, C., Wang, J., Ke, X., Feng, X., Yang, Y., Ren, H., Huang, Z. (2023). Experimental study on working mechanics of PDC bit with worn teeth, *Diamond & Abrasives Engineering*, Vol. 43, No. 1, 35-42.
- [10] Yang, Y., Xie, S., Cai, C., Cao, Y., Li, B. (2023). A study on the mechanism of heat generation law of PDC wear tooth cutting, *Journal of Southwest Petroleum University (Science & Technology Edition)*, Vol. 45, No. 1, 180-188, [doi: 10.11885/j.issn.1674-5086.2020.06.04.03](https://doi.org/10.11885/j.issn.1674-5086.2020.06.04.03).
- [11] Liu, W., Deng, H., Liu, Y., Chen, X., He, C., Zhu, X. (2023). Experimental investigation of the rock cutting process with blunt PDC cutters, *Geoenergy Science and Engineering*, Vol. 226, Article No. 211803, [doi: 10.1016/j.geoen.2023.211803](https://doi.org/10.1016/j.geoen.2023.211803).
- [12] Baker Hughes. StabilisX shaped-cutter technology, from <https://www.bakerhughes.com/drilling/drill-bits/shapedcutter-technology/stabilisx-shapedcutter-technology>, accessed July 28, 2023.
- [13] Shao, F., Liu, W., Gao, D. (2021). Effects of the chamfer and materials on performance of PDC cutters, *Journal of Petroleum Science and Engineering*, Vol. 205, Article No. 108887, [doi: 10.1016/j.petrol.2021.108887](https://doi.org/10.1016/j.petrol.2021.108887).
- [14] Fu, Z., Tergeist, M., Kueck, A., Ostermeyer, G.-P. (2022). Investigation of the cutting force response to a PDC cutter in rock using the discrete element method, *Journal of Petroleum Science and Engineering*, Vol. 213, Article No. 110330, [doi: 10.1016/j.petrol.2022.110330](https://doi.org/10.1016/j.petrol.2022.110330).

- [15] Akbari, B., Miska, S. (2016). The effects of chamfer and back rake angle on PDC cutters friction, *Journal of Natural Gas Science and Engineering*, Vol. 35, Part A, 347-353, doi: [10.1016/j.jngse.2016.08.043](https://doi.org/10.1016/j.jngse.2016.08.043).
- [16] Jamaludin, A.A., Mehat, N.M., Kamaruddin, S. (2019). Optimizing wear rate for PDC cutter using Taguchi's technique and response surface methodology a comparative analysis, *International Journal of Engineering & Technology*, Vol. 8, No. 3, 155-161.
- [17] Detournay, E., Defourny, P. (1992). A phenomenological model for the drilling action of drag bits, *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, Vol. 29, No. 1, 13-23, doi: [10.1016/0148-9062\(92\)91041-3](https://doi.org/10.1016/0148-9062(92)91041-3).
- [18] Zhang, Z., Zhou, Q., Zhang, K., Zheng, G., Zhang, T. (2019). Analysis of the influence of rock cutting depth on the temperature distribution of cutter, *Journal of China Coal Society*, Vol. 44, No. S2, 492-501, doi: [10.13225/j.cnki.jccs.2019.0949](https://doi.org/10.13225/j.cnki.jccs.2019.0949).
- [19] Ju, P., Tian, D., Tian, H. (2023). Simulation and experimental study on rock disintegration characteristics of special-shaped PDC cutters, *Gazi University Journal of Science*, Vol. 36, No. 1, 414-428, doi: [10.35378/guis.927956](https://doi.org/10.35378/guis.927956).
- [20] Yahiaoui, M., Paris, J.-Y., Delbé, K., Denape, J., Gerbaud, L., Dourfaye, A. (2016). Independent analyses of cutting and friction forces applied on a single polycrystalline diamond compact cutter, *International Journal of Rock Mechanics & Mining Sciences*, Vol. 85, 20-26, doi: [10.1016/j.ijrmms.2016.03.002](https://doi.org/10.1016/j.ijrmms.2016.03.002).
- [21] Munoz, H., Taheri, A., Chanada, E. (2016). Rock cutting characteristics on soft-to-hard rocks under different cutter inclinations, *International Journal of Rock Mechanics and Mining Sciences*, Vol. 87, 85-89, doi: [10.1016/j.ijrmms.2016.05.014](https://doi.org/10.1016/j.ijrmms.2016.05.014).
- [22] Rostamsowlat, I., Richard, T., Evans, B. (2018). An experimental study of the effect of back rake angle in rock cutting, *International Journal of Rock Mechanics and Mining Sciences*, Vol. 107, 224-232, doi: [10.1016/j.ijrmms.2018.04.046](https://doi.org/10.1016/j.ijrmms.2018.04.046).
- [23] Zhu, X.H., Wang, Y.F., Liu, W.J., Tan, B., Luo, Y.X., Li, Z.L. (2023). A new method for evaluating the rock cutting and breaking performance of PDC cutters in heterogeneous granites, *Natural Gas Industry*, Vol. 43, No. 4, 137-147.
- [24] Ju, P., Tian, D., Wang, C., Tian, H. (2021). Theoretical and simulation analysis on rock breaking mechanical properties of arc-shaped PDC bit, *Energy Reports*, Vol. 7, 6690-6699, doi: [10.1016/j.egy.2021.09.148](https://doi.org/10.1016/j.egy.2021.09.148).
- [25] Mazen, A.Z., Rahmanian, N., Mujtaba, I.M., Hassanpour, A. (2021). Effective mechanical specific energy: A new approach for evaluating PDC bit performance and cutters wear, *Journal of Petroleum Science and Engineering*, Vol. 196, Article No. 108030, doi: [10.1016/j.petrol.2020.108030](https://doi.org/10.1016/j.petrol.2020.108030).
- [26] Xiong, C., Huang, Z.-W., Wang, L.-C., Shi, H.-Z., He, W.-H., Chen, Z.-L., Li, G.-S. (2023). Study on rock breaking characteristics and mechanism of conical PDC cutter, *Rock and Soil Mechanics*, Vol. 44, No. 8, 2432-2444.
- [27] Kenneth, E., Russell, S.C. (2016). Innovative ability to change drilling responses of a PDC bit at the rigsite using interchangeable depth-of-cut control features, In: *Proceedings of the IADC/SPE Drilling Conference and Exhibition*, Fort Worth, Texas, USA, 1-10, doi: [10.2118/178808-MS](https://doi.org/10.2118/178808-MS).
- [28] Wang, Y.L., Yang, L., Chen, J.H., Li, P. (2023). Supply chain game analysis based on mean-variance and price risk aversion under different power structures, *Advances in Production Engineering & Management*, Vol. 18, No. 1, 104-115, doi: [10.14743/apem2023.1.460](https://doi.org/10.14743/apem2023.1.460).
- [29] Ma, Y.C., Zhang, P., Huang, Z.Q., Deng, R., Yu, H.M. (2020). Optimal design for cutter-layout of global force balanced PDC bits, *China Mechanical Engineering*, Vol. 31, No. 20, 2412-2419, doi: [10.3969/j.issn.1004-132X.2020.20.003](https://doi.org/10.3969/j.issn.1004-132X.2020.20.003).
- [30] Liu, S.B., Ni, H.J., Wang, Y., Zhang, H. (2021). Mechanism of multi-dimensional impact loads applied in increasing the rock cutting efficiency of a PDC bit, *Journal of Vibration and Shock*, Vol. 40, No. 2, 258-264, doi: [10.13465/j.cnki.jvs.2021.02.035](https://doi.org/10.13465/j.cnki.jvs.2021.02.035).
- [31] Fakhimi, A., Lanari, M. (2014). DEM-SPH simulation of rock blasting, *Computers and Geotechnics*, Vol. 55, 158-164, doi: [10.1016/j.compgeo.2013.08.008](https://doi.org/10.1016/j.compgeo.2013.08.008).
- [32] Das, R., Cleary, P.W. (2010). Effect of rock shapes on brittle fracture using Smoothed Particle Hydrodynamics, *Theoretical and Applied Fracture Mechanics*, Vol. 53, No. 1, 47-60, doi: [10.1016/j.tafmec.2009.12.004](https://doi.org/10.1016/j.tafmec.2009.12.004).
- [33] Bennett, K.C., Regueiro, R.A., Luscher, D.J. (2019). Anisotropic finite hyper-elastoplasticity of geomaterials with Drucker-Prager/Cap type constitutive model formulation, *International Journal of Plasticity*, Vol. 123, 224-250, doi: [10.1016/j.ijplas.2018.11.010](https://doi.org/10.1016/j.ijplas.2018.11.010).
- [34] Huang, Z., Ma, Y., Li, Q., Xie, D. (2017). Geometry and force modelling, and mechanical properties study of polycrystalline diamond compact bit under wearing condition based on numerical analysis, *Advances in Mechanical Engineering*, Vol. 9, No. 6, 1-15, doi: [10.1177/1687814017702080](https://doi.org/10.1177/1687814017702080).

Dynamic price competition market for retailers in the context of consumer learning behavior and supplier competition: Machine learning-enhanced agent-based modeling and simulation

Deng, G.F.^{a,*}

^aNational Taichung University of Science and Technology, College of Languages, Applied Information & Japanese Program, Taichung, Taiwan (R.O.C.)

ABSTRACT

This study analyzes the impact of consumer learning behavior and supplier price competition on retailer price competition in a complex adaptive system. Using machine Learning-enhanced agent-based modeling and simulation, the study applies fuzzy logic and genetic algorithms to model price decisions, and reinforcement learning and swarm intelligence to model consumer behavior. Simulations reveal that different learning behaviors result in different retailer competition patterns, and that supplier price competition affects the strength of retailer price competition. Simulation results demonstrate that consumer learning behavior influences retailer competition, with self-learning consumers leading to higher-priced partnerships, and collective-learning consumers leading to a shift in price competition among retailers. In contrast, perfect rationality consumers result in low-price competition and the lowest average margin and profit. Additionally, the competitive price behavior of suppliers impacts retailers' price competition patterns, with supplier price competition reducing retailer price competition in the perfect rationality consumer market and enhancing it in the self-learning and collective-learning consumer markets, leading to lower average prices and profits for retailers. This study presents a simulated market for price competition among suppliers, retailers, and consumers that can be expanded by subsequent scholars to test related hypotheses.

ARTICLE INFO

Keywords:

Pricing competitive model;
Complex adaptive system (CAS);
Agent-based modeling and simulation (ABMS);
Machine learning (ML);
Genetic algorithms (GA);
Fuzzy logic (FL);
Reinforcement learning (RL);
Swarm intelligence (SW);
Consumer learning behavior

*Corresponding author:

raymaldeng@gmail.com
(Deng, G.F.)

Article history:

Received 10 April 2023
Revised 14 November 2023
Accepted 16 November 2023



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

1. Introduction

Competitive markets are often complex because they are the emergent result of many individual agents (e.g., consumers, retailers, supplier) whose motivations and actions combine so that even simple behavioral rules can result in surprising patterns [1, 2]. Agent-based modeling and simulation (ABMS) is a rich platform for studying complex evolving systems to test behavioral economics theory and bridge micro and macro models. Many studies are beginning to apply ABMS to investigate how individual agents interact during the competitive process and achieve a balanced outcome from the perspective of a dynamic evolutionary game [3-8]. ABMS frameworks

are applied to design artificial adaptive agents to simulate the decision-making and learning behaviors of real-world individuals and their interactions with incomplete information and partially rational decisions. In this way, the payoffs and final equilibrium outcomes of individuals can be observed over time to support and assist marketing strategies [9]. The advantage of ABMS is that it can produce results similar to those of the analytic model under the same assumptions and can be further extended and relaxed to analyze dynamic simulation results in more complex situations, test and develop theories, and provide strategic implications [10-14].

The earliest assumptions of the price-competitive model were to explore the equilibrium of the final price-competitive situation when oligopolistic competitors sell homogeneous products without regard to their capacity, when they have the same cost function, when demand is certain and known, and when consumers choose only the lower-priced products [15]. Hunt's general theory of competition describes the complexity and evolutionary nature of competitive markets from a different perspective [16]. The theoretical view is that competition in the market is caused by market imbalances. Competitive behavior stems from the endogenous learning behavior of suppliers. Tay and Lusch applied ABMS to construct a producer competitive market and observed the price competitive dynamics and equilibrium of suppliers. The results validate Hunt's general theory of competition [17].

The subsequent development of the price competition model includes a number of considerations, such as consumer different cognitive decisions [18], heterogeneous consumers with objective supplier preferences [19] and switching behavior [20], consumers' sensitivity to price information [21], consumers' demonstrated loyalty [22], consumers' social network word-of-mouth [23], consumers' different levels of learning ability [24], consumers' ability to demonstrate strategic purchasing behavior etc. [25]. In summary, previous studies that have examined the price competitive market from a dynamic perspective have focused on price competition among suppliers and incorporated consumer learning behavior to construct an ABMS market to observe the dynamic effects of learning behavior on overall market prices but have neglected the role of retailers in the price competitive market [26].

In a price competitive market characterized by independent suppliers and retailers, the dynamics of price setting and adjustment are influenced by three main competitive forces. First, there is the competition that occurs at the supplier level, as each supplier seeks to offer the most attractive price for the product they produce relative to other suppliers in the market. Second, there is the competition that occurs at the retailer level, where retailers seek to set the most attractive prices for the set of goods they offer, taking into account the prices set by their competitors. Third, there is the vertical interaction competition that occurs between the suppliers and retailers, where the two parties negotiate and adjust prices in response to each other's decisions and actions. Together, these three forces shape the competitive landscape of the market and ultimately influence the price decisions of the firms involved [15, 27]. The price decisions of retailers nowadays have a significant degree of influence on the market. The impact of retailer competition on the price competitive market is an issue worth clarifying [14, 28].

In this study, ABMS is applied to construct a price competition simulation market involving retailers, suppliers, and consumers, where each retailer, supplier, and consumer can establish their behavioral decision rules as adaptive agents. Individuals will interact with each other to maximize their rewards and be given learning behaviors to observe over time in a complex adaptive system (CAS) simulation to understand the intricate competitive relationships between retailers, suppliers, and consumers [29].

Evolutionary game theory has been applied to price competition, emulating the behavior of human deductive reasoning and inductive reasoning, where suppliers and retailers have price decision rules and learned behavior mechanisms that apply fuzzy logic (FL) to their price decisions and genetic algorithms (GA) to the price rules adjustment [27, 30]. In constructing consumer purchase learning behavior, reinforcement learning (RL) from psychology was used to model self-purchase learning behavior [12, 30], while swarm intelligence from biology was used to model collective learning consumers [32].

This study will answer the question: from a price competition perspective, is it possible to observe how different *consumer learning behavior* and *supplier price competition behavior* affect

the price competition process of retailers and the eventual price equilibrium outcome that may result? This study will examine the following:

- The impact of three different *consumer learning behaviors*, namely *perfect rationality*, *self-learning*, and *collective learning* on retailers' dynamic co-opetition strategies.
- The impact of the *non-price competitive* and *price competitive* behavior of suppliers on retailers' dynamic co-opetition strategies.

2. Pricing competition model

The price competitive market in this study comprises a number of suppliers, retailers, and consumers. The suppliers produce a single product, and their price decisions are primarily setting the wholesale price of that product. The suppliers' products are homogeneous, and there is price competition between suppliers. Retailers are responsible for selling the products produced by their respective suppliers. The retailer's pricing decisions primarily determine the margins of the supplier's products, which are also price-competitive with each other. The supplier's wholesale price and the retailer's margins are added together to form the retail price. There is still a so-called price competition between suppliers and retailers. In this price-competitive interaction, suppliers and retailers make price decisions with the aim of maximizing their own profit. In addition, they follow human deductive reasoning and inductive reasoning behavior, apply fuzzy logic as the basis for their price decisions and genetic algorithms as a way of adjusting price rules through empirical learning, resulting in an evolutionary game of price competition between retailers, between suppliers, and between suppliers and retailers.

In competitive markets, consumers engage in purchase learning behavior. In a market, the price of each retailer is a form of incomplete information to the consumer. As a result, consumers judge which retailer they can get the best price from based on their past purchasing experience. Once in the retailer's shop, the consumer can directly compare prices between suppliers and choose the lower price. Consumers can use the rewards of this purchase to form an experience and adjust their choice of retailer for the next time through learned behavior. In addition, consumers can model self-learning behavior about purchases through reinforcement learning in psychology [12] and the learning behavior of a group of consumers through swarm intelligence in biology [32].

Price competition means that suppliers will consider the wholesale prices of other suppliers and margins of downstream retailers when making price decisions. Retailers will consider the margins of other retailers and wholesale prices of upstream suppliers when making price decisions. *No price competition* means that there is peaceful co-existence *between suppliers* and *between suppliers and retailers*. As a result, suppliers' price decisions remain fixed throughout the simulation period. At the same time, the retailer's price decisions are made without regard to the supplier's price.

2.1 Price competitive market

The price competitive market in this study consists of supplier S_i ($i = 1, \dots, i^*$) that each produces its own supplier's product and has the same fixed cost c_i^f . Retailer R_j ($j = 1, \dots, j^*$) sells each supplier S_i 's product with the same fixed cost c_j^f . At each point in time t , supplier S_i determines the wholesale price w_i for the product. The w_i set for the period applies to all retailers R_j ($j = 1, \dots, j^*$). Next, retailer R_j determines the margin m_j for that product. The m_j set for the period applies to all suppliers S_i ($i = 1, \dots, i^*$). The sum of m_j and w_{ij} determines the retail price of S_i at R_j , R_{ij} . In the market, there are many heterogeneous consumers C_k ($k = 1, \dots, k^*$). Consumers have different price sensitivities, forgetting rate, degree of rationality and different propensities towards retailers $\theta_{k,j}$ ($j = 1, \dots, j^*$). The profit of S_i ($i = 1, \dots, i^*$) and R_j ($j = 1, \dots, j^*$), as well as the accumulated capital $profit_i$, $profit_j$, can be calculated based on the demand q_{ij} for retailer R_j 's product S_i generated by the consumer's purchase behavior.

The price competitive market simulation works as follows: At each point in time t , the interactive steps described below are included:

- Step 1: In the mode, the simulation time is t^* , which contains multiple time points t . Each time point t contains multiple rounds r^* . Each round r contains multiple encounters ε^* .
- Step 2: For each round r , the supplier and retailer select a price decision rule from a library of price decision rules to conduct an ε^* encounters test.
- Step 3: For each round ε , the supplier refers to the previous price interaction with other competing suppliers and retailers and applies fuzzy logic to determine its own wholesale price, and proposes its wholesale price to each retailer.
- Step 4: The retailer takes into account its margin with other competing retailers and the supplier's wholesale price in the previous round and applies fuzzy logic to determine its own margin.
- Step 5: Consumers will be able to evaluate and choose a retailer based on their purchasing decisions. The consumer does not have access to the current prices set by the retailer for each supplier's product but must enter the retailer's shop in order to obtain price information. After deciding on a retailer, the consumer can directly choose the supplier with the lower price, which in turn generates the demand q_{ij} for the supplier M_i among the retailers R_j and the profit of each retailer and supplier in that transaction.
- Step 6: Consumers use purchase learning behavior to adjust their purchase decisions based on the rewards generated by this epsilon purchase decision. The learning behavior is modified using self- and collective learning.
- Step 7: When $\varepsilon < \varepsilon^*$, return to step 3 and $\varepsilon = \varepsilon + 1$. Otherwise, determine if $r < r^*$ is valid, and if it is, go back to step 2 with $r = r + 1$, and if it is not, go to step 8.
- Step 8: Retailers and suppliers evaluate the performance of their price decisions based on profit and apply genetic algorithms to adjust their price decisions.
- Step 9: Determine whether t reaches the maximum simulation time t^* , and stop if it does. Otherwise, go back to step 2, $t = t + 1$.

2.2 Price competitive behavior of supplier and retailer

The study explores the modeling of human deductive and inductive reasoning through the continuous adjustment of pricing strategies by suppliers and retailers using fuzzy decision rules, with the ultimate goal of enhancing their survival. The FL-GA theoretical framework provides an excellent opportunity to showcase the effectiveness of ABMS in this context. A schematic representation of the FL-GA architecture is illustrated in [17]. The process of fuzzy rules adjustment for improved decision-making involves two methods: exploitation, which involves the recombination of existing genetic material in novel ways via crossover, and exploration, which involves the adoption of new genetic material via mutation. Selection and reproduction are used to keep and replicate successful decision-making approaches while discarding those that are ineffective. Both exploitation and exploration strategies rely on the use of selection and reproduction to optimize decision-making performance.

These rules are constantly evaluated and adjusted based on their predictive accuracy in forecasting market behavior. Successful rules are retained and acted upon while poorly performing rules are discarded. The discarded rules are replaced with new hybrid rules generated from the effective ones. Additionally, new rules are generated and tested in response to emerging market information. This iterative process of learning and adaptation allows suppliers and retailers to continually adapt to the constantly evolving market conditions. Please refer to the study by Tay and Lusch for details on the FL-GA theoretical framework [17].

2.3 Consumer learning behavior

Each consumer is an artificial adaptive agent and will exhibit bounded rationality learning behavior in the face of incomplete information about prices. After each purchase, the individual will calculate the price difference based on his or her past purchase experience or observation of other people's purchases as a reward for the purchase and adjust his or her propensity to buy

from each retailer. The aim is to get the same product at a lower price the next time. The study distinguishes three types of purchase learning behavior as follows:

Type I *Perfect Rationality*: This type was designed as a control group. The consumer is provided with all price information prior to purchase. There is no price information search cost for the consumer and the lowest price for each purchase, so there is no learning behavior, and it can be used as a basis for comparison with other learning behaviors.

Type II *Self-Learning*: Because this type of consumer has the highest price search costs, they will only adjust their purchase decisions after each purchase by comparing them to their last purchase price. The consumer decides which retailer to go to this time to get a better price based on his or her past purchase experience and decides which retailer to go to this time. The self-learning behavior is based on reinforcement learning (RL) algorithms.

Type III *Collective-Learning*: Collective-Learning emphasizes a way of learning by comparing purchasing experiences with those of others as a basis for future revision decisions. Consumers use the swarm intelligence algorithm to compare prices between groups after making a purchase decision to see if the purchase is more expensive or cheaper, and then adjust their propensity for the retailer to facilitate a higher return on their next purchase decision.

2.3.1 Self learning behavior

RL means that actions that produced good results in the past will be reinforced, making them more likely to be taken again in the future. Actions that produced bad results in the past will be weakened, making them less likely to be taken again in the future. Action means that the consumer decides which retailer to buy from. The result is a *price differential* in the price paid for the product purchased. With incomplete information, the current price is only known after the consumer has chosen the retailer, and the price information for other retailers is still not available and must be recalled from past memory. When the return on price differential is perceptually positive, the likelihood of choosing that retailer next time increases, and vice versa. Such a learning model is based entirely on one's own experience and fully expresses the spirit of self-adaptive learning. Therefore, RL algorithms were used to model this self-learning behavior. The relevant behavioral patterns and algorithm flow are as follows.

Step 1: Initial action

RL is the conversion of past experience into a propensity to act on that decision. The level of propensity is seen as a preference for a particular *Action*. A consumer's propensity is his or her preference for a retailer θ_j . For example, a consumer who chooses between two retailers and whose purchase action is either to shop at retailer R_1 or to shop at retailer R_2 , is given an $\theta_{j=1}$ and $\theta_{j=2}$ to indicate the consumer's propensity for each of the two actions. Parameters are set for each consumer, including their initial propensity to shop at the retailer, upper and lower limits of propensity, rationality, price sensitivity, etc.

Step 2: Rules of selection

According to the rules of selection for retailers, the *rules of selection* are used to determine the retailers that consumers decide to buy from at this stage, based on their current propensity for each retailer. The higher an individual's propensity for a particular action, the more likely that individual is to choose that action. The choice of retailer is determined by the ratio of each retailer's propensity to the total propensity. This ratio indicates the probability of each retailer being selected. The formula for the *choice of action* rule is as follows:

$$x_j(t) = \begin{cases} \max \theta_j(t) & \text{if } rand < \beta \\ \frac{\theta_j(t)}{\sum_{j=1}^n \theta_j(t)} & \text{else} \end{cases} \quad (1)$$

The variable $x_j(t)$ is the consumer's choice of retailer j at time point t , and β is the degree of rationality of the consumer's purchase decision. A higher β value indicates a higher probability of choosing a retailer with a higher propensity to buy. Random variable $rand$: between 0 and 1. When $rand < \beta$, the consumer directly chooses the retailer with the highest propensity; otherwise, the consumer is given a chance to choose the retailer with the highest propensity according to the ratio of each retailer's propensity and then directly chooses the supplier with the lowest price.

Step 3: Calculation of reward

The price at which the consumer purchases at this point in time t is compared to the price of the previous purchase ($t - 1$) to calculate the return on the consumer's choice of retailer for this purchase. The formula for calculating the reward at time point t is as follows.

$$u_{k,j}^t = \alpha_k (p_j^{t-1} - p_{k,j}^t) \tag{2}$$

The variables are defined as follows. $u_{k,j}^t$ is the reward from retailer j arising from consumer k 's choice to buy from retailer j . α_k is price sensitivity of the consumer k . The greater the price sensitivity, the greater the effect on the price differential. $P_{k,j}^t$ is the price at which consumer k buys at this time at retailer j . p_j^{t-1} is the price of the consumer k 's last purchase. $p_j^{t-1} - p_{k,j}^t$ is price differential.

Step 4: The propensity to update decisions

The reward generated by a consumer action is the key factor that allows for the *reinforcement* of learning. Each time a decision is made, RL updates the reward to a propensity for that strategy. The remaining strategies that are not selected are not updated because of the learning effect. Past propensities are partially lost over time. Therefore, through repeated purchases, each propensity will increase or decrease after each purchase, making it more or less likely to be chosen next time. The propensity update method for each retailer is:

$$\begin{aligned} \theta_s(t+1) &= (1 - \delta)\theta_s(t) + u_s(t) \\ \theta_u(t + 1) &= (1 - \delta)\theta_u(t) \end{aligned} \tag{3}$$

$\theta_s(t)$ indicates the propensity to update for the selected retailer s and $\theta_u(t)$ is the propensity to update for the unselected retailer u . $1 \geq \delta \geq 0$ is a memory parameter (recency), which indicates that past experiences or memories are forgotten over time. A larger δ value indicates a greater emphasis on the most recent memory. Over time, consumers will have a relatively high propensity to buy from retailers that generate high returns. The probability of being selected is relatively high. In the end, the consumer's action set tends to be simple, and learning tends to be stable.

2.3.2 Collective learning behavior

Collective learning is generally conceptualized as a dynamic and cumulative process that results in the production of knowledge. Such knowledge is institutionalized in the form of structures, rules, routines, norms, discourse, and strategies that guide future action. This study uses Swarm Intelligence to model consumer collective-learning behavior. Swarm Intelligence primarily mimics the process of a bird's collective flight in search of a food location. Each individual modifies the intensity of the flight by taking into account both their own past experience and the experience of other individuals in the flock to determine the direction of the next flight.

In collective purchasing learning behavior, consumers also decide which retailer to buy from based on the rules of retailer choice. Under the retailer choice rule, consumers decide which retailer to choose at this stage based on their current propensity for each retailer through the *choice action rule*. The higher an individual's propensity for a particular action, the more likely that individual is to choose that action. The choice of retailer is determined by the ratio of each retailer's propensity to the total propensity. The ratio indicates how likely each retailer is to be selected, where the rules of selection are the same as for self-learning. After each purchase, con-

sumers adjust their propensity to buy from a particular retailer by calculating rewards based on their past purchasing experience and the experience of others, as expressed in the following equation:

$$\begin{aligned} u_{k,j}^{t+1} &= \alpha_k \times \text{rand}() [(p_{k,j}^{t-1} - p_{k,j}^t) + (p_g^t - p_{k,j}^t)] \\ \theta_{k,j}^{t+1} &= \theta_{k,j}^t + u_{k,j}^{t+1} \end{aligned} \quad (4)$$

$u_{k,j}^{t+1}$ is the reward generated by the consumer k 's choice of retailer i , which can be used as a basis for adjusting the consumer's propensity for each retailer. α_k is the price sensitivity of consumer k . rand is The degree of randomness indicates an extraneous environmental variable, a random number between 0 and 1. $p_{k,j}^{t-1}$ is the retailer price at which consumer k purchased in the previous purchase. $p_{k,j}^t$ is the retailer price at which consumer k purchased this purchase. p_g^t is the retailer whose price is the lowest among the group at time point t . $\theta_{k,j}^t$ is consumer k ' propensity for various retailers j at this stage. $p_g^t - p_{k,j}^t$ is consumer perception of the price difference.

3. Experimental results

3.1 Setup for price competitive market simulation

In this study, the Matlab programming language was used to implement the ABMS system. The parameters of the simulation were set as shown in the table 1. The number of retailers n_r is set to 2, the number of suppliers n_m is set to 2, and the number of consumers n_c is set to 100. The simulation was conducted 25 times in each market environment setting. Simulation time $t^* = 1000$. A time point t consists of 10 rounds (r^*). Each retailer competes on 4 encounters (ϵ^*) per round, and the average margin, average profit, and the cumulative profit of the two retailers are recorded for each time point t . Finally, the average of the 25 experiments is taken to produce the experimental data. This study found that equilibrium was reached after 1000 simulation time t^* . Retailers' price competition patterns are repeated. It was therefore decided that 1000 simulation time t^* would be the time for the simulation.

Table 1 Parameter settings for the price competitive market

	Initial Settings for the Variables/Parameters	Setup Values/Range
Retailer/ Supplier	Number of retailers n_r / Number of suppliers n_m	2/2
	Fixed costs (c_i & c_j)	400
	Initial assets	1000
	Initial wholesale price ($w_i^{t=0}$)	8
	Initial margin $m_{i,j}^r(t=0)$	8
	Upper and lower price limits	6-10
	Number of rules for the library (n_{rule})	16
	Number of semantic values of input variables under fuzzy rules	4
	Number of semantic values of output variables under fuzzy rules	8
	Mating rate	0.8
	Mutation rate	0.2
	Code	Binary encoding
	Fitness function	Total sales
	Mating operator	Two-point mating
	Consumer	Number of individual consumers (n_c)
Initial retailer propensity of consumers (θ_j): Normal distribution		$\tilde{\mu} = 8 \quad \sigma = 0.5$
Upper bound of consumer propensity to retailer		10
Lower bound of consumer propensity to retailer		2
Price sensitivity (α)		1
Degree of rationality (β)		0.8
System settings	Consumer forgetting rate (δ)	0.2
	System implementation language	Matlab
	Simulation time (t^*)	1000
	Number of consumer purchases per cycle (r^*)	40
	Number of experiments per market environment	25 times

The retailer determines the unit margin for the product. The unit margin is summed with the wholesale price of the product and becomes the selling price. The unit margin is set at a maximum of \$10 and a minimum of \$6 per unit. The retailer's fixed cost per operation is (c_i) \$400. The initial unit profit $m_{i,j}^r(t=0)$ is set at \$8. In a non-competitive situation, the wholesale price is set at a fixed price of \$8 per unit of product. In the case of a competitive supplier, the wholesale price is determined using fuzzy logic. The wholesale price per unit is also set at a maximum of \$10 and a minimum of \$6 per unit. The supplier's fixed cost (c_i) per production run is \$400. The initial wholesale price $w_i(t=0)$ is set at \$8. The initial propensity of consumers toward retailers presents θ_i^k a normal distribution, where $\tilde{\mu} = 8$ and $\sigma = 0.5$. The lower bound of consumer propensity towards retailers ranges from 2 to 10. Consumer price sensitivity α is designed to be 1, and consumer forgetting rate δ is 0.2. The degree of rationality of consumers' purchase decisions was set at 0.8.

3.2 Impact of price competition among retailers

In this study, three different consumer learning behaviors and two different suppliers competitive behaviors were simulated to obtain a total of six market environment settings. Table 2 presents that the different *consumer learning behavior* and *supplier price competition behavior* affect the overall average margin, the overall cumulative profit, and the profit gap at the retailer end.

First, we looked at the impact of three different *consumer learning behaviors* on retailers' competition relationship. Market 1 shows that the overall average margin (6.7842) and cumulative profit (-2,845,861) at the retailer end are the lowest when consumers demonstrate *perfect rationality* and suppliers do not engage in price competition. It can be inferred that retailers are more likely to make higher profits through low-price strategies when consumers have immediate access to price information, therefore, more likely to engage in low-price competition among retailers. As a result of low-price competition, the average cumulative profit of retailers is also the lowest.

If we further analyze the price competition pattern in Fig. 1, we can see that during the price competition between the two retailers, margin tend to fall quickly to the lowest price (\$6) under low price competition. After that, both parties have the opportunity to see the benefits of cooperation, and prices are gradually raised to \$8, but only for a relatively short period of time (50 rounds). The retailer learns that it is more rewarding to set lower price than its rival. As a result, the retailer will begin to compete at a lower price. The cycle goes on and on, with no equilibrium of low-price convergence. The validity of the basic price competition model can be verified by the results of the experiment in which consumers exhibit *perfect rationality* behavior, which is the same as the Bertrand model of competition, as mentioned in the previous literature. Under the same assumptions, it can be found that price competition among producers eventually approaches the equilibrium price of the lowest price.

Market 2 shows that the overall average margin price (\$9.7216) and cumulative profit (\$3,435,891) at the retailer end are highest when consumers exhibit *self-learning behavior* and suppliers do not engage in *price-competitive behavior*. According to the company, consumers who do not have immediate access to price information must rely on their past purchasing experience to make purchasing decisions. As a result, retailers set prices high for long periods of time and only occasionally run low-price promotions to retain consumers and reduce price competition among retailers.

Further analysis of the price competition pattern in Fig. 2 shows that in the early stages of price competition between the two retailers, those who adopt a low-price strategy are able to make higher profits. The retailer with a high-price strategy continues to lose money but does not compete at lower prices. After a certain period, the retailer with the low-price strategy finds that higher prices will lead to higher profits and gradually raises its prices. Thereafter, the two retailers positioned their prices at a high level. Even when prices are adjusted, only sporadic low-price fluctuations occur.

Table 2 Experimental results

Market environment	Consumer learning behavior	Competitive behavior of suppliers	Overall average margin on the retailer side	Overall cumulative profit on the retailer side	Overall cumulative profit gap on the retailer side	Chart coding
1	Perfect rationality	N/A	6.7842	-2,845,861	1,932,295	Fig. 1
2	Self-learning	N/A	9.7216	3,435,891	3,686,571	Fig. 2
3	Collective-learning	N/A	9.6558	3,282,325	2,776,079	Fig. 3
4	Perfect rationality	Yes	7.1787	-2,060,160	4,451,600	Fig. 4
5	Self-learning	Yes	9.3794	2,633,278	7,582,923	Fig. 5
6	Collective-learning	Yes	9.3208	2,472,801	6,780,293	Fig. 6

Market 3 shows that with consumers exhibiting *collective-learning behavior* and suppliers not engaging in *price-competitive behavior*, the overall average margin (\$9.6558) and cumulative profit (\$3,435,891) at the retailer end is slightly lower than *self-learning behavior* but not significantly different. This is a more interesting result. From a practical point of view, if consumers adjust their purchasing propensity for retailers by making inter-group comparisons after making a purchase, this should result in more intense competition among retailers for lower prices. Further analysis in Fig. 3 reveals that the difference in competitive patterns between Markets 2 and 3 is that: Whilst retailers' prices are positioned at a high level, and there is sporadic and very short-lived low-price competition (2-5 rounds) between retailers in self-learning, collective-learning is likely to experience more frequent and short-lived medium-price competition (lasting 20-50 rounds). However, the modeling results do not suggest that long-term low-price competition will occur.

Comparing Fig. 2 and Fig. 3 on the comparison of consumer propensity and cumulative profit patterns at the retailer end, it can be seen that there is little difference between the propensity of consumers who can engage in self-learning behavior towards the two retailers and the change in propensity is relatively gentle. This means that consumers show a more gradual adjustment when their propensity for retailers changes. In contrast, consumers who are able to engage in collective-learning have a high frequency of change and a steeper curve of change in their propensity for retailers. This means that consumers tend to adjust their propensity for both retailers in a wide range of ways, immediately, quickly, and frequently. Then we looked at the average profit pattern of retailers per round. The self-learning behavior presented shows that retailers take longer (more rounds) to make a lead transition per round. Collective-learning shows that retailers often have short lead transitions. This is due to post-facto group price comparison by consumers. Price information can still be spread over a short period of time. A one-time low-price strategy can sometimes have an effect.

Next, we looked at the impact of the *non-price competitive* and *price competitive* behavior of suppliers on retailers' co-opetition relationship. The study found that there were similarities with the *non-competitive behavior* of suppliers. For example, the retailer with the lowest average price and cumulative profit (\$7.18) was found in the case of *perfect rationality* behavior by consumers. Retailers presented an overall average margin price (\$9.35 vs. \$9.32) and cumulative profit (\$2,633,278 vs. \$2,472,801) that were fairly similar under self-learning and collective-learning, with retailers under self-learning still being slightly higher than retailers under collective-learning.

Market 4 shows that overall average margin and cumulative profit at the retailer end are lower than the overall average margin and overall profit, given the *perfect rationality* of consumers and the *competitive pricing practices* of suppliers. Looking further at Markets 1 vs 4, it can be seen that *price competition* by suppliers increases the overall average margin and overall cumulative profit at the retailer end. However, when comparing Markets 2 vs 5 and Markets 6 vs 3, it can be seen that the overall average margin price and cumulative profit at the retailer end decreases when suppliers engage in *price competition*. The reason for this is inferred to be that if

there is price competition from suppliers, this will result in more drastic price changes in the market, which in turn will require longer purchase learning behavior on the part of consumers to recognize the pattern of price competition among retailers. As a result, retailers would need to increase the learning rewards for consumers by offering more low prices.

If we further analyze Figs. 1 and 4, we can see that if suppliers are competing on price, the price competition pattern does not drop in the short term and remain locked at the lowest price (\$6) for a long period of time. Instead, the price will fluctuate in a relatively gentle cycle between \$6 ~\$8. This pattern is also reflected in the profit patterns of retailers in each round. Price competition from suppliers has resulted in less lead shifting and wider lead gap for retailers.

If we further analyze Figs. 2 and 5, we can see that if suppliers are competing on price, the price competition pattern does not rise in the short term and remain locked at the higher price (\$10) for a long period of time. Instead, the price will fluctuate in a relatively gentle cycle between \$9 ~\$10. In terms of consumer propensity and retailers' profit patterns in each round, consumer propensity to swap leadership becomes more frequent, the gap becomes wider, and the speed of adjustment in propensity increases if suppliers engage in price competition. Retailers' profit patterns show an increase in the likelihood of mutual and significant profit leads, which is unlike the high price levels over time, fewer lead swaps, and less of a profit gap seen in Market 2.

If we further analyze Figs. 3 and 6, we can see that if there is competition from suppliers, the price competition pattern does not produce the same short-term increase in price competition and long-term lock-in at high prices that would occur if suppliers did not compete. Instead, there would be several short-term price competitions at prices falling between \$8.50 and \$9.

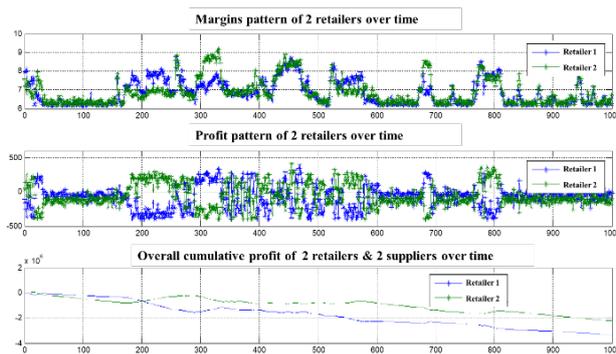


Fig. 1 Market 1: Non-competitive suppliers & perfect rationality consumers

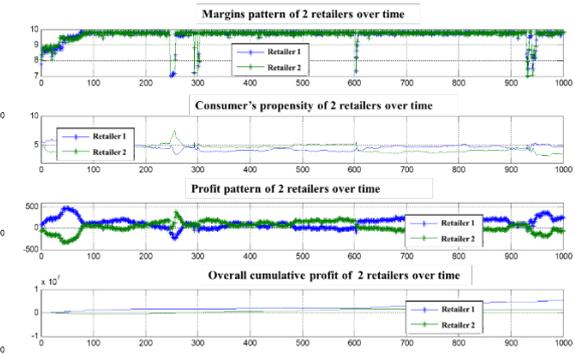


Fig. 2 Market 2: Non-competitive suppliers & self-learning behavior consumers

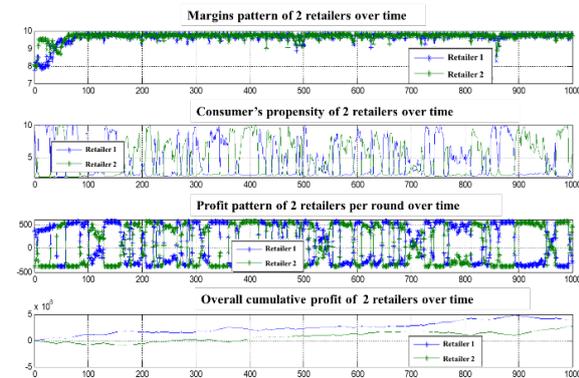


Fig. 3 Market 3: Non-competitive suppliers & collective-learning consumers

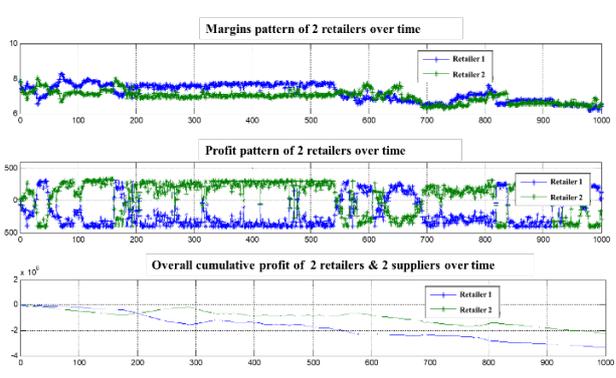


Fig. 4 Market 4: Competitive suppliers & perfect rationality consumers

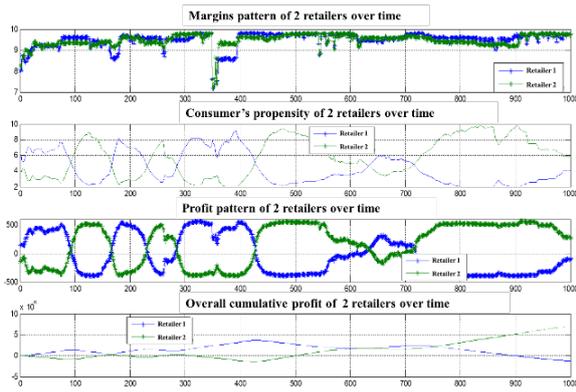


Fig. 5 Market 5: Competitive suppliers & self-learning consumers

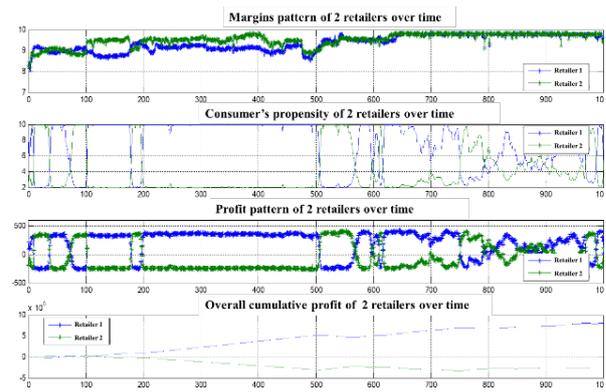


Fig. 6 Market 6: Competitive suppliers & collective-learning consumers

In terms of consumer propensity and the profit pattern of retailers in each round, the propensity and profit of retailers in Market 3 changes most sharply in the absence of competitive behavior by suppliers. The adjustment in propensity is steep, and there is a high degree of frequent switching of leadership. Where suppliers compete, the change in propensity and profit tends to moderate. Retailers with a profitable lead remain steadily and significantly ahead for a period of time, but lead changes can still occur in very short periods of time.

Overall, there is competition for the lowest price at the retailer's end as consumers display *perfect rationality* behavior. In the case of self-learning behavior, if suppliers do not compete on price, retailers will cooperate, and all adopt a high price strategy. The overall average price and overall cumulative profit at the retailer end will be the highest, with the least amount of interchange between retailers and the least change in profit. However, if suppliers were to compete on price, this would increase the intensity of market change. In the absence of price competition by suppliers, the most dramatic changes in retailer propensity and profit patterns occur in the absence of Collective-Learning behavior, although retailer price competition can result in high prices, including large leads, the most frequent lead switching, and the steepest lead transitions. Price competition by suppliers can help to mitigate the intensity of market change.

4. Conclusions

The contribution of this study is to present a simulated market for price competition involving suppliers, retailers, and consumers and observe how different 'consumer learning behavior' and *supplier price competition behavior* affect the price competition process of retailers and the eventual price equilibrium outcome.

For the retailer side, this study found that consumers engaged in self-learning behaviors that were most beneficial to the retailer side of the competition. Collective-learning has the potential to become a major learning method at a time when e-commerce and online communities are becoming more prevalent. The impact of this could also bring the overall profit of the retailer side closer to that of the consumer market where self-learning is present. It is important to note, however, that if consumers engage in collective-learning, this will lead to increased competition for the retailer as a whole. Therefore, individual retailers need to be cautious in dealing with collective-learning consumers and be aware of market dynamics and be flexible in adjusting their pricing strategies to avoid becoming loss-making retailers due to rigid pricing strategies.

In the face of price competition from suppliers, only in the control group (i.e., the *perfect rationality* consumer market) can the overall retailer side see an increase in profit. Other environments will result in lower profit. Therefore, in a market where consumers are likely to *self-learning* and *collective-learning*, it is more advantageous for retailers as a whole to stabilize the wholesale price of suppliers. From a retailer's perspective, where suppliers do not compete, if consumers engage in self-learning behavior, this will result in the highest overall average price for the retailer and maximize overall cumulative profit. Therefore, from a practical point of view,

a retailer can achieve higher profits if it can keep supplier prices stable and induce consumers to engage mainly in self-learning behavior.

The price competitive market proposed in this study also models the price competitive behavior of retailers under the basic assumptions of the Bertrand model (no competition among suppliers, *perfect rationality* among consumers). The experimental results are not only identical to the Bertrand model of competition but also show the dynamic process of price competition. In the further, the price competitive market can be extended or expanded by subsequent scholars to include specific factors to observe the results of the interaction and test related hypothesis.

Acknowledgement

The authors are grateful to the MOE Teaching Practice Research Program , ROC, for the partial financial support of this work under the grant PGE1120918 .

References

- [1] Cioffi-Revilla, C. (2010). A methodology for complex social simulations, *Journal of Artificial Societies and Social Simulation*, Vol. 13, No. 1, Article No. 7, doi: [10.18564/jasss.1528](https://doi.org/10.18564/jasss.1528).
- [2] Dominguez, R., Cannella, S. (2020). Insights on multi-agent systems applications for supply chain management, *Sustainability*, Vol. 12, No. 5, Article No. 1935, doi: [10.3390/su12051935](https://doi.org/10.3390/su12051935).
- [3] Izquierdo, L.R., Izquierdo, S.S., Sandholm, W.H. (2019). An introduction to ABED: Agent-based simulation of evolutionary game dynamics, *Games and Economic Behavior*, Vol. 118, 434-462, doi: [10.1016/j.geb.2019.09.014](https://doi.org/10.1016/j.geb.2019.09.014).
- [4] Zhang, R., Sun, B. (2019). A competitive dynamics perspective on evolutionary game theory, agent-based modeling, and innovation in high-tech firms, *Management Decision*, Vol. 58, No. 5, 948-966, doi: [10.1108/md-06-2018-0666](https://doi.org/10.1108/md-06-2018-0666).
- [5] Pathak, S.D., Dilts, D.M., Biswas, G. (2007). On the evolutionary dynamics of supply network topologies, *IEEE Transactions on Engineering Management*, Vol. 54, No. 4, 662-672, doi: [10.1109/tem.2007.906856](https://doi.org/10.1109/tem.2007.906856).
- [6] Xiao, M., Tian, Z.Y. (2022). Evolutionary game analysis of company collaborative strategy in cloud manufacturing platform environment, *Advances in Production Engineering & Management*, Vol. 17, No. 3, 295-310, doi: [10.14743/apem2022.3.437](https://doi.org/10.14743/apem2022.3.437).
- [7] Bojic, S., Maslaric, M., Mircetic, D., Nikolicic, S., Todorovic, V. (2023). Simulation and Genetic Algorithm-based approach for multi-objective optimization of production planning: A case study in industry, *Advances in Production Engineering & Management*, Vol. 18, No. 2, 250-262, doi: [10.14743/apem2023.2.471](https://doi.org/10.14743/apem2023.2.471).
- [8] Wang, Y.L., Yang, L., Chen, J.H., Li, P. (2023). Supply chain game analysis based on mean-variance and price risk aversion under different power structures, *Advances in Production Engineering & Management*, Vol. 18, No. 1, 104-115, doi: [10.14743/apem2023.1.460](https://doi.org/10.14743/apem2023.1.460).
- [9] Tay, N.S.P., Lusch, R.F. (2007). Agent-based modeling of ambidextrous organizations: Virtualizing competitive strategy, *IEEE Intelligent Systems*, Vol. 22, No. 5, 50-57, doi: [10.1109/MIS.2007.4338494](https://doi.org/10.1109/MIS.2007.4338494).
- [10] Vilà, X. (2008). A model-to-model analysis of Bertrand competition, *Journal of Artificial Societies and Social Simulation*, Vol. 11, No. 2, 1-11.
- [11] Zhang, T., Zhang, D. (2007). Agent-based simulation of consumer purchase decision-making and the decoy effect, *Journal of Business Research*, Vol. 60, No. 8, 912-922, doi: [10.1016/j.jbusres.2007.02.006](https://doi.org/10.1016/j.jbusres.2007.02.006).
- [12] Hopkins, E. (2007). Adaptive learning models of consumer behavior, *Journal of Economic Behavior & Organization*, Vol. 64, No. 3-4, 348-368, doi: [10.1016/j.jebo.2006.02.010](https://doi.org/10.1016/j.jebo.2006.02.010).
- [13] Siebers, P.-O., Aickelin, U., Celia, H., Clegg, C.W. (2010). Simulating customer experience and word-of-mouth in retail-a case study, *Simulation*, Vol. 86, No. 1, 5-30, doi: [10.1177/0037549708101575](https://doi.org/10.1177/0037549708101575).
- [14] Rand, W., Rust, R.T. (2011). Agent-based modeling in marketing: Guidelines for rigor, *International Journal of Research in Marketing*, Vol. 28, No. 3, 181-193, doi: [10.1016/j.ijresmar.2011.04.002](https://doi.org/10.1016/j.ijresmar.2011.04.002).
- [15] Choi, S.C. (1996). Price competition in a duopoly common retailer channel, *Journal of Retailing*, Vol. 72, No. 2, 117-134, doi: [10.1016/S0022-4359\(96\)90010-X](https://doi.org/10.1016/S0022-4359(96)90010-X).
- [16] Grassl, W. (2003). A general theory of competition: Resources, competences, productivity, economic growth by Shelby D. Hunt, *The Review of Austrian Economics*, Vol. 16, 385-393, doi: [10.1023/A:1027353307279](https://doi.org/10.1023/A:1027353307279).
- [17] Tay, N.S.P., Lusch, R.F. (2005). A preliminary test of Hunt's general theory of competition: Using artificial adaptive agents to study complex and ill-defined environments, *Journal of Business Research*, Vol. 58, No. 9, 1155-1168, doi: [10.1016/j.jbusres.2004.04.005](https://doi.org/10.1016/j.jbusres.2004.04.005).
- [18] Dulam, R., Furuta, K., Kanno, T. (2021). Quantitative decision-making model to analyze the post-disaster consumer behavior, *International Journal of Disaster Risk Reduction*, Vol. 61, Article No. 102329, doi: [10.1016/j.ijdrr.2021.102329](https://doi.org/10.1016/j.ijdrr.2021.102329).
- [19] Stummer, C., Kiesling, E., Gunther, M., Vetschera, R. (2015). Innovation diffusion of repeat purchase products in a competitive market: An agent-based simulation approach, *European Journal of Operational Research*, Vol. 245, No. 1, 157-167, doi: [10.1016/j.ejor.2015.03.008](https://doi.org/10.1016/j.ejor.2015.03.008).

- [20] Du, H., Jiang, Y. (2019). Backup or reliability improvement strategy for a manufacturer facing heterogeneous consumers in a dynamic supply chain, *IEEE Access*, Vol. 7, 50419-50430, doi: [10.1109/access.2019.2911620](https://doi.org/10.1109/access.2019.2911620).
- [21] Zhang, W.-G., Zhang, Q., Mizgier, K.J., Zhang, Y. (2017). Integrating the customers' perceived risks and benefits into the triple-channel retailing, *International Journal of Production Research*, Vol. 55, No. 22, 6676-6690, doi: [10.1080/00207543.2017.1336679](https://doi.org/10.1080/00207543.2017.1336679).
- [22] Wei, Y., Li, F. (2015). Impact of heterogeneous consumers on pricing decisions under dual-channel competition, *Mathematical Problems in Engineering*, Vol. 2015, Article ID 584084, doi: [10.1155/2015/584084](https://doi.org/10.1155/2015/584084).
- [23] Zadbood, A., Hoffenson, S. (2022). Social network word-of-mouth integrated into agent-based design for market systems modeling, *Journal of Mechanical Design*, Vol. 144, No. 7, Article No. 071701, doi: [10.1115/1.4053684](https://doi.org/10.1115/1.4053684).
- [24] Avitabile, D., Hoyle, R., Samaey, G. (2014). Noise reduction in coarse bifurcation analysis of stochastic agent-based models: An example of consumer lock-in, *SIAM Journal on Applied Dynamical Systems*, Vol. 13, No. 4, 1583-1619, doi: [10.1137/140962188](https://doi.org/10.1137/140962188).
- [25] Du, H., Xiao, T. (2019). Pricing strategies for competing adaptive retailers facing complex consumer behavior: Agent-based model, *International Journal of Information Technology & Decision Making*, Vol. 18, No. 6, 1909-1939, doi: [10.1142/s021962201950038x](https://doi.org/10.1142/s021962201950038x).
- [26] Bentley, P.J., Lim, S.L. (2022). From evolutionary ecosystem simulations to computational models of human behavior, *Wires Cognitive Science*, Vol. 13, No. 6, Article No. e1622, doi: [10.1002/wcs.1622](https://doi.org/10.1002/wcs.1622).
- [27] Jiang, L., Wang, Y. (2009). Supplier competition in decentralized assembly systems with price-sensitive and uncertain demand, *Manufacturing & Service Operations Management*, Vol. 12, No. 1, 93-101, doi: [10.1287/msom.1090.0259](https://doi.org/10.1287/msom.1090.0259).
- [28] Negahban, A., Yilmaz, L. (2014). Agent-based simulation applications in marketing research: An integrated review, *Journal of Simulation*, Vol. 8, No. 2, 129-142, doi: [10.1057/jos.2013.21](https://doi.org/10.1057/jos.2013.21).
- [29] Vargo, S.L., Koskela-Huotari, K., Baron, S., Edvardsson, B., Reynoso, J., Colurcio, M. (2017). A systems perspective on markets – Toward a research agenda, *Journal of Business Research*, Vol. 79, 260-268, doi: [10.1016/j.jbusres.2017.03.011](https://doi.org/10.1016/j.jbusres.2017.03.011).
- [30] Negahban, A., Giabbanelli, P.J. (2022). Hybrid agent-based simulation of adoption behavior and social interactions: Alternatives, opportunities, and pitfalls, *IEEE Transactions on Computational Social Systems*, Vol. 9, No. 3, 770-780, doi: [10.1109/TCSS.2021.3101794](https://doi.org/10.1109/TCSS.2021.3101794).
- [31] Bozdoğan, A., Görkemli Aykut, L. Demirel, N. (2023) An agent-based modeling framework for the design of a dynamic closed-loop supply chain network, *Complex & Intelligent Systems*, Vol. 9, 247-265, doi: [10.1007/s40747-022-00780-z](https://doi.org/10.1007/s40747-022-00780-z).
- [32] Zhang, T., Brorsen, B. (2009). Particle swarm optimization algorithm for agent-based artificial markets, *Computational Economics*, Vol. 34, No. 4, 399-417, doi: [10.1007/s10614-009-9171-9](https://doi.org/10.1007/s10614-009-9171-9).

IoT-based Deep Learning Neural Network (DLNN) algorithm for voltage stability control and monitoring of solar power generation

Shweta, R.^{a,*}, Sivagnanam, S.^b, Kumar, K.A.^c

^aAnnamalai University, Tamilnadu, India

^bAnnamalai University, Tamilnadu, India

^cBHEL, Tamilnadu, India

ABSTRACT

Today, Solar Photovoltaic (SPV) energy, an advancing and attractive clean technology with zero carbon emissions, is widely used. It is crucial to pay serious attention to the maintenance and application of Solar Power Generation (SPG) to harness it effectively. The design was more costly, and the automatic monitoring is not precise. The main objective of the work related to designed and built up the Internet of Things (IoT) platform to monitor the SPV Power Plants (SPVPP) to solve the issue. IoT platform designing and Data Analytics (DA) are the two phases of the proposed methodology. For building the IoT device in the IoT platform designing phase, diverse lower-cost sensors with higher end-to-end delivery ratio, higher network lifetime, throughput, residual energy, and better energy consumption are considered. Then, Sigfox communication technology is employed at the Low-Power Wireless Area Network (LPWAN) communication layer for lower-cost communication. Therefore, in the DA phase, the sensor monitored values are evaluated. In the analysis phase, which is the most significant part of the work, the input data are first pre-processed to avoid errors. Next, to monitor the Energy Loss (EL), the fault, and Potential Energy (PE), the solar features are extracted as of the pre-processed data. The significance of utilizing the Transformation Search centered Seagull Optimization (TSSO) algorithm, the significant features are chosen as of the extracted features. Therefore, the computational time of the solar monitoring has been decreased by the Feature Selection (FS). Next, the features are input into the Gaussian Kernelized Deep Learning Neural Network (GKDLNN) algorithm, which predicts the faults, PE, and EL. In the experimental evaluation, solar generation is assessed based on Wind Speed (WS), temperature, time, and Global Solar Radiation (GSR). The systems are satisfactory and produce more power during the time interval from 12:00 PM to 1:00 PM. The performance of the proposed method is evaluated based on performance metrics and compared with existing research techniques. When compared to these techniques, the proposed framework achieves superior results with improved precision, accuracy, F-measure, and recall.

ARTICLE INFO

Keywords:

Solar photovoltaic (SPV);
Internet of things (IoT);
Data analytics;
Sigfox communication technology;
Low-power wireless area network (LPWAN);
Energy loss;
Machine learning;
Transformation search centered seagull optimization algorithm (TSSO);
Gaussian kernelized deep learning Neural Network (GKDLNN)

*Corresponding author:

shwetaraj434@gmail.com
(Shweta, R.)

Article history:

Received 3 April 2023
Revised 13 December 2023
Accepted 15 December 2023



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

1. Introduction

Owing to the economic and sustainable characteristics, the RE is drawing considerable attention in research [1]. Solar energy has several benefits despite more RE being available. Here, via a Photovoltaic System (PVS), electrical energy is acquired. The solar panels which can directly convert sunlight or solar energy into electricity are included in this system [2]. Temperature,

irradiance, voltage, and current are the fundamental parameters that affect the solar panel's efficiency. Owing to the disparities in the temperature, solar irradiance, weather conditions, together with several other factors, the power produced from SPV installations is vulnerable [3]. Therefore, a real-time solar monitoring system is necessary for augmenting the PV panel's performance by contrasting it with the experimental outcome to start preventative action [4, 5]. The PV Monitoring System (PVMS), which is based on both wired and wireless networks, is designed to transmit parameters to a remote coordinator. This coordinator provides a web-based application for remote access [6]. Recently, the SPVMS has been incorporated with a wireless platform that involves data acquisition as of diverse sensors along with nodes via wireless data transmission [7, 8]. In the project data's practical verification and in the optimization of power plants' conversion efficacy, data acquisition is extremely significant [9-12]. The IoT facilitates objects to be recognized along with managed remotely over the prevailing system, developing new doors for the physical world's unadulterated integration into computer-centered systems and resulting in enhanced accuracy, economic benefits, together with productivity [13, 14]. For transmitting the data to the cloud, sensors and microcontrollers play an extremely vital part in relation to Wi-Fi module [14, 15-17]. Amassing data, a network layer for the data's transmission, a data processing layer for processing essential data, and finally, an application layer that operates as an interface betwixt end devices and the network all are the four-layer IoT structure in the solar-centered IoT monitoring system [18]. Wireless sensor networks, GPS, 2G/3G/4G, GPRS, GSM, WI-FI, microcontroller, RFID, microprocessor, etc. accomplish this IoT [19, 20]. On a range of parameters like energy extracted, energy potential, historical generation analysis, fault identification, and related ELs, the PVMS's intention is to offer accurate data [21]. Tracking the panel voltage, temperature, current, and the solar system's real data is to be synchronized as of time to time are the numerous difficulties faced by the solar PVS's remote monitoring [22]. There is an assortment of advantages in utilizing IoT, like enhanced accuracy, efficacy, lesser human involvement, along with cost reduction, which is exhibited by numerous investigations [23, 24]. On the monitor, the historical data and extracted energies are effortlessly presented. The manual computation takes a longer time when faults, PE, and ELs are not manually computed. Therefore, for SPVPP's DA, this work exhibits the design together with the development of an IoT platform.

2. Literature review

Kandimalla and Kishore [25] designed a prototype for the implementation of a cost-efficient technique based on IoT. This prototype checks an SPVPP for performance assessment using open-source tools and resources such as Thingspeak and Arduino. Thingspeak, a Software as a Service (SaaS) platform, provides web space for monitoring parameters. On Website designing and maintenance, Thingspeak offered every service for free of cost, which saved many investments. Thus, for the ordinary people who set up rooftop solar plants, it concentrated on a lower-cost system with a simple interface so that they could monitor the solar plants effortlessly, devoid of depending on service providers. Besides real-time monitoring, it facilitated Fault Detection (FD), preventive maintenance, along with plant's historical analysis.

Almonacid-Olleros *et al.* [26] examined the output power generation's evaluation by human-crafted characteristics with numerous temporal windows along with Deep Learning (DL) methodologies to acquire relative outcomes concerning the PV system's analytical models regarding error metrics along with learning time. In a PVS with IoT capacities created within the Opera Digital Platform underneath the Univer Project, the surrounding data along with ground truth of energy production had amassed, which was installed for two decades on the Campus of the University of Jaén (Spain). When analogized to the existing analytical model, the machine learning models provided enhanced outcomes, with considerable dissimilarities in learning time together with performance. For enriching the performance, the multiple temporal windows' utilization had exhibited as an appropriate tool to model temporal features.

Ramamurthi and Nadar [27] concentrated on an SPP's interdisciplinary field in India's southern region. It was integrated with the latest digital technology of IoT that bolstered to check real-time data of temperature, solar irradiation, panel voltage, peak power, current, together with the

tilt angle of solar panels. Also, to manage the solar panel's optimum inclination angle for augmenting PV power utilizing the Genetic Algorithm (GA). For enhancing the system's tracking behaviour, an integrated SPP with IoT had constructed with hardware setup along with it was virtually tested. The performance of monitoring together with maintenance of power plant parameters with cost-efficiency had considerably improved by the IoT-centered control of SPPs [28-31].

Shapsough *et al.* [32] proffered an IoT-centered architecture, which employs IoT hardware, software, along with communication technologies to facilitate real-time monitoring together with the management of SPVMS at huge scales. This system facilitated stakeholders to remotely regulate and observe the PVSs along with to inspect the diverse environmental factor's effects like soiling, air quality, along with the weather. Regarding the network delay along with resource consumption, the system was executed and examined. For wide-ranging real-time communication, Message Queueing Telemetry Transport (MQTT) was utilized. Since the average network delay was lesser than 1 s, it was verified that the architecture was perfect for solar along with smart grid monitoring systems. The assessment exhibited that the hardware consumes about 3 % of the panel's output for resource consumption, whilst the application as well used an extremely small percentage of the CPU [33]. Since the lower-cost constrained edge devices deployed the architecture in an excellent way where the incorporation of IoT-centered paradigm, effective MQTT communication, along with lower-resources consumption made this system cost-efficient and scalable.

Pulungan *et al.* [34] explained the design, manufacture along with testing of a single axis online solar tracker monitoring system with the IoT. In real-time, the monitoring outcome's data retrieval occurred and was exhibited in the graphic data's form. Next, the data acquired as of sensors was linked to a microcontroller, and then linked to a WIFI module [8]. Samkria *et al.* [30] presented IoT along with LabVIEW-centered automatic FD of 3×3 solar array systems for controlling and observing internet connectivity distantly [35-36]. In the PVS, to produce a panel alert for damaged panels, the handling of the GUI indicator aided the monitoring system. Through internet connectivity, node MCU in the receiver segment facilitated fault status' transmission of PV arrays. Intended for visualizing the 3×3 PV array's fault status, the IoT-centered Blynk app was utilized. The Blynk's dashboard visualized each array with the status.

3. Methodology

3.1 Design and development of IOT platform for analytics of solar power

A larger percentage of the energy produced as of renewable resources is shared by the SPV, which is one of the well-known sustainable energy sources. Monitoring technologies have attained substantial focus corresponding to the performance advancement since the necessity for solar energy has increased immensely in recent decades. Presently, via wireless data transmission, the SPV-MS has been incorporated with the wireless platform, which includes data acquisition as of several sensor nodes. Nevertheless, signal interference, larger data management, security, and long-range data transmission are the challenges that influence the performance of SPV-MS. Therefore, with the DA of S-PVPP monitoring, the IoT platform has been developed here regarding various sensor devices along with varied communication technologies. IoT platform designing along with DA of SPV data was encompassed in this proposed methodology. Sensors, solar panel model, Wi-Fi module, and authentication are explicated in the IoT platform design. In DA, pre-processing, feature extraction, FS, and analysis steps are performed to assess the data. By wielding the TSSO model, the FS is conducted. Then, the GKDLNN is employed for the assessment. Fig. 1 exhibits the proposed methodology's block diagram.

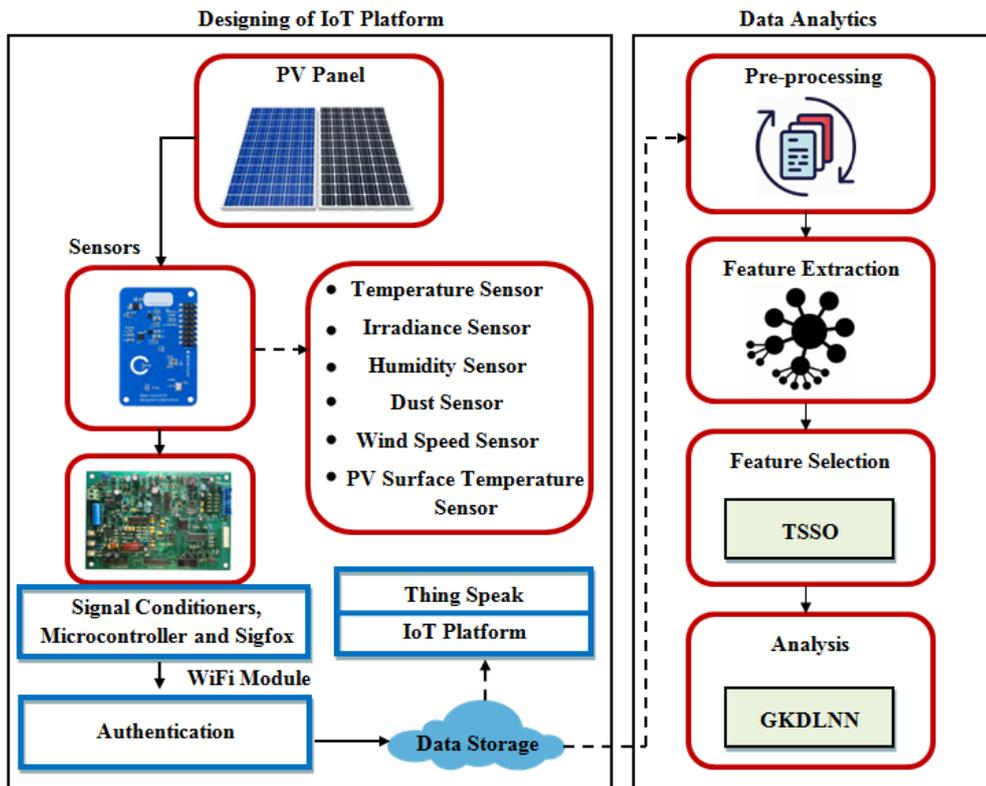


Fig. 1 A Block diagram for the proposed research methodology

3.2 Designing IoT platform

Initially, to obtain the SPV data, the IoT platform is designed. Firstly, the SPV model is structured for the design. Next, to sense the values with higher end-to-end delivery ratio, lower cost, higher network lifetime, residual energy, better energy consumption, throughput, along with better performance, the sensor of IoT devices is chosen. Then, the authentication phase is executed. The device is permitted to forward the data if it is an authorized device. The Wi-Fi module was structured to transfer the data. In this, the Sigfox technology at the LPWAN communication layer is employed for communication purposes. A greater performance was offered by the Sigfox technology at a lower cost.

3.3 Solar PV model

The sunlight is converted into DC electrical energy with a set of parallel along with series PV cells, which are included in the PV panels. At the PV modules' terminals, the observation of the characteristics desires the accumulation of extra parameters to the basic equation:

$$H = H_{PV} - H_o \left[\exp \left(\frac{L+R_s H}{\lambda L_t} \right) - 1 \right] - \frac{K+R_s H}{R_p} \tag{1}$$

The PV and saturation currents are specified as H_{PV} and H_o , the diode's ideality factor is signified as λ , the series and parallels resistances are defined as R_s and R_p , and the thermal voltage is proffered as L_t . The H_{PV} is expressed as:

$$H_{PV} = (H_{PV,n} + M_l \Delta E) \frac{Q}{Q_n} \tag{2}$$

The light generated current is notated as $H_{PV,n}$, the short circuit current-temperature coefficient is proffered as M_l , the temperature difference betwixt the actual and nominal values is illustrated as E , the solar irradiation and its nominal value are described as Q and Q_n . $H_{PV,n}$ in conjunction with H_o is formulated as:

$$H_{PV,n} = \frac{R_p + R_s}{R_p} H_{sc,n} \tag{3}$$

$$H_o = \frac{H_{sc,n} + M_l \Delta E}{\exp\left(\frac{L_{oc,n} + M_v \Delta E}{\lambda L_t}\right) - 1} \quad (4)$$

The short circuit current in conjunction with open-circuit voltages underneath the nominal condition is represented as $H_{sc,n}$, $L_{oc,n}$, the open-circuit voltage per temperature coefficient is indicated as M_v . To generate the desired output power, the larger PVPP contains numerous PV modules, which are linked in series-parallel. This PVPP is modelled as:

$$H = O_p H_{PV} - O_p H_o \left[\exp\left(\frac{L + R_s \left(\frac{O_m}{O_p}\right) H}{O_m \lambda L_t}\right) - 1 \right] - \frac{L + R_s \left(\frac{O_m}{O_p}\right) H}{R_p \left(\frac{O_m}{O_p}\right)} \quad (5)$$

The number of series-connected modules in a string is specified as O_m ; in addition, the number of parallel-connected strings is signified as O_p . The DC-DC Buck-boost converter is welded in the SPV module; similarly, to maximize the power, the Maximum Power Point Tracking (MPPT) is utilized.

3.4 Measurement of sensors

Various sensors are utilized to gauge the PV model's characteristics. In this, six varied sensors are considered. The solar irradiance SP110, ambient temperature DS18B20, dust GP2Y1010AU0F, humidity DHT22, WS anemometer, and the PV module surface temperature sensor PT100 are read by the six sensors. The terms in the sensor device SD are denoted as:

$$SD = \{Te, Ir, Hu, Du, Wi, St\} \quad (6)$$

The temperature sensor device is specified as Te , the irradiance device is signified as Ir , the humidity device is illustrated as Hu , the dust device is proffered as Du , the wind device is defined as Wi , and the surface temperature sensor device is symbolized as St .

Temperature sensor: The DS18B20 sensor is utilized to gauge the temperature. The DS18B20 is a 1-wire programmable Temperature sensor. With a decent accuracy of ± 5 °C, an extensive range of temperature can be measured from -55 °C to +125 °C. The values in the hard along with soft environments are gauged by this sensor.

Irradiance sensor: The SP110 sensor, which is a self-powered, analog sensor with a 0-400 mV output, is employed to measure solar irradiance. An effortless connection of dataloggers along with controllers is achieved by this sensor, which integrates a silicon-cell photodiode with a rugged, self-cleaning sensor housing design, together with a higher-quality cable terminating in a pre-tinned pigtail.

Humidity sensor: The DHT22 sensor, which is an essential, lower-cost digital temperature along with a humidity sensor, is employed for gauging the humidity. To gauge the surrounding air, a capacitive humidity together with a thermistor is utilized; subsequently, it discharges a digital signal on the data pin (analog input pins are not required). It has several benefits like having 0-100 % humidity readings with 2-5 % accuracy. In addition, it is better for -40 °C to 80 °C temperature readings with ± 0.5 °C accuracy.

Dust sensor: A simple air monitoring module with onboard Sharp GP2Y1010AU0F is termed a Dust Sensor. Fine particles, which are greater than 0.8 μm in diameter, even like cigarette smoke, are discovered by this sensor. The sensor's analog voltage output is linear with dust density. To sustain an extensive range of power supplies, an embedded voltage boost circuit has been utilized by this module.

Wind speed sensor: For utilizing this module, the black wire is linked to the power along with the signal ground, the brown wire is connected to 7-24VDC (9 V is utilized with success); in addition, the analog voltage is measured on the blue wire. The voltage ranges from 0.4 V (0 m/s wind) up to 2.0 V (for 32.4 m/s WS).

PV surface temperature sensor: The PT100 sensor is employed for gauging the PV surface temperature. Therefore, the sensor is steady, fast, along with precise. The IP65 field, which is an onsite 2-point calibration, has exceptional long-term stability.

3.5 Authentication

The sensor nodes must be registered on the server to transfer the data. The data as of the authorized sensor devices alone is received by the server. Initially, the sensor devices are registered. Details about sensor devices are provided in the registration phase. Therefore, in the database, the details are amassed. The device login to the server at the time device along with personal password is provided to transmit the data. In the verification phase, the device is permitted to broadcast the data to the server only if the details presented in the login phase are existed in the database or else the device is specified as an unauthorized device. The authentication factors are signified as:

$$Rg \rightarrow (DN, pwd, GC) \quad (7)$$

$$Ln \rightarrow (DN, pwd) \quad (8)$$

$$Vr \rightarrow \begin{cases} \text{Allowed, if } Ln = Rg \\ \text{Not allowed, otherwise} \end{cases} \quad (9)$$

The authentication's registration, verification, and verification process are notated as Rg , Ln and Vr , the sensor device's device name and general characteristics are symbolized as DN and GC , and the password is exhibited as pwd .

Wi-Fi Module: Prior to amassing the computed data on an IoT server or cloud, it is processed by the AT Mega 328 (ESP8266) utilizing a Wi-Fi module. The renowned Thingspeak IoT platform is used for regular, weekly, and monthly data analysis.

Data storage: The sensed data are amassed in the cloud server with the aid of the Wi-Fi module. Petabytes of data, which are required to be amassed, processed, along with evaluated, are produced by sensor networks. In storage services, an effectual role is performed by cloud computing.

Data Analytics: The data being gathered are evaluated here. On the monitor, the energy along with historical power generation is displayed. The automatic monitoring process is desired since the prediction fault, EL, and PE are not conducted effortlessly in the manual process. The procedures in automatic DA are explicated below.

Pre-processing: Firstly, the pre-processing is performed. Some missing values along with repeated values might occur in the sensor data. An error is caused at the end of the evaluation if there occurs any missing along with repeated value in the dataset. The mean value of before together with after values of the missing place is computed to replace the missing values. The mean computation is expressed as:

$$T_i = \frac{T_{i+1} + T_{i-1}}{2} \quad (10)$$

The after value as of the missing value is represented as T_{i+1} and the before value is indicated as T_{i-1} . By employing the Hadoop-MapReduce (MR) function, the repeated data are eliminated. For writing applications, Hadoop-MR, which is a programming along with software framework, is utilized. Here, an enormous amount of data is processed speedily on larger clusters of computed nodes. Map phase and reduce phase are the two phases utilized here. Every single phase contains key-value pairs as input as well as output. Regarding the data structured (key, value) in pairs, the map along with reduce functions of MR are proffered. A set of input key-value pairs is taken during computation; thus, producing a set of output key-value pairs. The map along with reduce functions in Hadoop-MR is formulated as:

$$M_p(\alpha_1, T_1) \rightarrow l_t(\alpha_2, T_2) \quad (11)$$

$$C_e(\alpha_2, l_t(T_2)) \rightarrow l_t(T_2) \quad (12)$$

The map function is specified as M_p , the key values are signified as α_1 and α_2 , the reduce function is defined as C_e and the list of MR functions is proffered as l_t .

3.6 Feature extraction

Subsequent to pre-processing, from the pre-processed data, the features are extracted. Relative Humidity, Irradiance, Temperature, PV surface Temperature, WS, and Dust Accumulation are the features being extracted. The feature terms are notated as:

$$Z_s = \{z_1, z_2, z_3, \dots, z_n\}, \text{ or } z_i, i = 1, 2, \dots, n \quad (13)$$

The feature set is indicated as Z_s and the n -number of features is depicted as z_n .

3.7 Feature selection

Here, to alleviate the monitoring process's computational time, as of the feature set being extracted, the significant features are chosen. The TSSO algorithm is deployed in the proposed model. The migration along with the behaviours of seagulls in nature is inspired for the development of the Seagull Optimization (SO) algorithm. Seagulls are a type of seabird. They encompass various species that primarily feed on fish, insects, amphibians, reptiles, and earthworms. This version is more direct and clearer. They are extremely intelligent and use their intelligence to find food and attack prey. For instance, they use breadcrumbs to attract fish, and to catch earthworms hidden underground, they mimic the sound of rain with their feet. A better performance was offered by the traditional SO algorithms. However, the issue of poor exploration along with exploitation searching capacity is faced them. Thus, in this paper, the transformation search is applied in the SO algorithm to address the aforementioned problem. The most crucial behaviours of seagulls are migration along with attack. The source food for seagulls is proffered as migration behaviours. The seagulls' attack towards the migrating birds at sea is mentioned as attack behaviour. Firstly, the populations are initialized. In this, the features being selected are regarded as the population. By the transformation search function, the population is initialized. It is represented as:

$$A_i = \varpi(lb + ub) - (z_i - \tilde{z}_i) \quad (14)$$

The initialized population set is proffered as A_i , the elastic factor is notated as ϖ , the upper and lower bound values are denoted as lb and ub . Next, the Fitness Value (FV) is computed. In this, the accurate evaluation is pondered as a FV. Following initialization, the migration along with the attacking behaviours is explicated individually.

(a) Migration behaviour

The algorithm imitates the seagulls' movement from one place to another during the migration process. Avoiding collisions, the best position's direction, and reaching the best position are the three criteria that must be satisfied by the seagulls. These criteria are explicated as follows. Initially, an extra variable F is presented to avoid collisions. Thus, betwixt the neighbours, the collision is prevented. For the new Search Agent (SA) position, the new variable is utilized.

$$LO_s = F W_s(t) \quad (15)$$

The SA's location that doesn't collide with other SAs is specified as LO_s , the SA's current location is indicated as W_s , the current iteration time is denoted as t , and the SA's mobile behaviour in the provided search space is signified as F .

$$F = \zeta_c - \left(t \left(\frac{\zeta_c}{ax_{ite}} \right) \right), t = 0, 1, 2, \dots, ax_{ite} \quad (16)$$

To regulate the utilization frequency of variable F , the parameter ζ_c is introduced. The variable F is linearly reduced from ζ_c to 0. After satisfying the condition of not colliding with other individuals, the seagull moves towards the best position. It is illustrated as:

$$SA_s = \delta (W_{bs}(t) - W_s(t)) \quad (17)$$

The position of current location Search Agent is indicated $SA W_s$ towards the best-fit $SA W_b$ (that is to say, the best seagull with a smaller FV) is represented as SA_s . The behaviour of δ , which is accountable for appropriate balancing betwixt exploration and exploitation, is randomized. It is gauged as:

$$\delta = 2 F^2 r \quad (18)$$

A random number which lies in the range of $[0, 1]$ is specified as r . Finally, regarding the best SA , the SA can update its position.

$$\varphi_s = |LO_s + SA_s| \quad (19)$$

The distance between the SA and best-fit SA (that is to say, best seagull whose FV is less) is denoted as φ_s .

(b) Attack behaviour

The seagulls make a spiral motion in the air whilst attacking the prey. This behaviours in u, v , and x planes are elucidated as:

$$u' = d \cos(y) \quad (20)$$

$$v' = d \sin(y) \quad (21)$$

$$x' = d y \quad (22)$$

$$d = p e^{yq} \quad (23)$$

The radius of every single turn of the spiral is specified as d , a random number in the range $[0 \leq y \leq 2\pi]$ is defined as y , constants to proffer the spiral shape are notated as p and q , and the natural logarithm's base is represented as e . The SA 's updated position is computed by wielding the Eqs. 15-23.

$$W_s(t) = (\varphi_s u' v' x') + W_{bs}(t) \quad (24)$$

The best solution is saved together with the position of other SA s is updated by the $W_s(t)$.

Input: Extracted features z_i

Output: Selected features \tilde{z}_i

Begin

Initialize population A_i , iteration t , and maximum iteration ax_i

Set $t = 1$

While ($t \leq ax_i$) **do**

Calculate Fitness function

//Migration Behaviour//

$r = \text{rand}(0,1)$ */* To generate the random number in range [0, 1] */*

$y = \text{rand}(0,2\pi)$ */* To generate the random number in range [0, 2\pi] */*

//Attacking Behaviour//

$d = p \times e^{yq}$ */* To generate the spiral behavior during migration */*

Calculate distance ϕ_i

$W_i(t) = (\phi_i \times u' \times v' \times x') + W_{bi}(t)$

Set $t = t + 1$

End while

Return optimal features \tilde{z}_i

End

Fig. 2 Pseudo code for TSSO

Fig. 2 depicts the TSSO's pseudo-code. Fig. 3 displays the flowchart for the TSSO.

The features being selected are modelled as:

$$\tilde{Z}_s = \{\tilde{z}_1, \tilde{z}_2, \tilde{z}_3, \dots, \tilde{z}_n\}, \text{ or } \tilde{z}_i, i = 1, 2, \dots, n \quad (25)$$

The feature set is notated as \tilde{Z}_s and the n -number of features is symbolized as \tilde{z}_n .

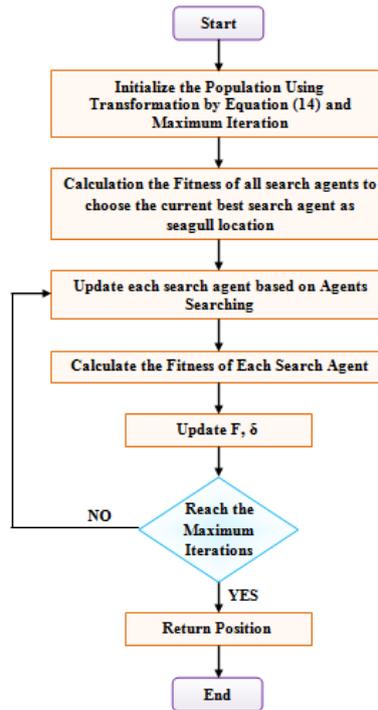


Fig 3 Flowchart for TSSO algorithm

3.8 DLLN analysis

Subsequent to FS, the features being selected are inputted to the analysis phase. Faults, PE, and EL are forecasted by this evaluation. In numerous applications, the Artificial Neural Network (ANN), which is an intelligent system, is utilized to address complex issues. Three layers are encompassed in the ANN structure. They are: (i) an Input Layer (IL), which includes the collected data, (ii) an Output Layer (OL), which generates computed data, along with (iii) one or more Hidden Layers (HLs), which are apt for linking the IL along with the OL. Collecting input and generating output are the two functions performed by a neuron, which is a fundamental processing unit of a NN. To obtain an output, every single input is multiplied by connection weights; the products and biases are added. Afterwards, they are passed via an Activation Function (AF). However, for the deep analysis, a single HL is possessed by the traditional NN. A more number of HLs is utilized in this proposed framework. A random weight initialization methodology is possessed by the baseline algorithm. In the output, a minor alteration in weight initialization can lead to a critical change, thereby worsening the outcome. Regarding the Gaussian neighbourhood function, the weight values are altered in the proposed methodology to conquer the aforementioned complication. Additionally, an effective AF is desired for the deep analysis. Consequently, the kernelized AF is utilized here. With this alteration, the NN is renamed as the GKDLNN algorithm. Fig. 4 displays the GKDLNN's structure.

To start with, the selected features \tilde{z}_i are inputted to the IL. Subsequently, the HL is provided with the IL's output. Regarding the weights, which is the sum of weighted synapse connections, the output is computed by the HL. The input is filtered by eliminating the redundant data and is forwarded to the subsequent HL for further processing. The HL process is mathematically formulated as:

$$hd_i = ib_s + \sum_{i=1}^n \tilde{z}_i \cdot \varepsilon_i \quad (26)$$

Here, the bias value is notated as ib_s , the weight value is symbolized as ε_i , and the input feature is signified as \tilde{z}_i . By employing the Gaussian neighbourhood function, the weight adjustment is derived as:

$$\varepsilon_{i+1} = \varepsilon_i + NN_i[\tilde{z}_i - \varepsilon_i] \quad (27)$$

The Gaussian neighbourhood function is modelled as NN_i .

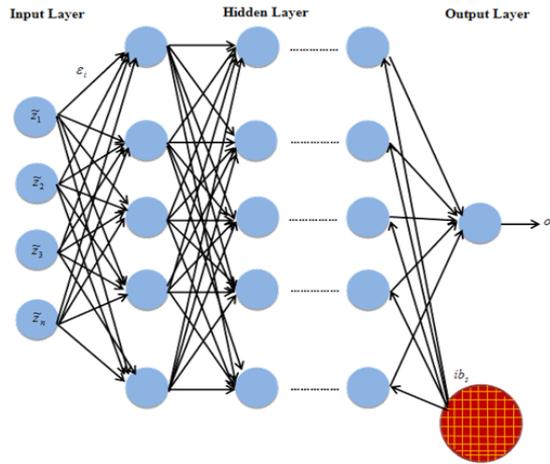


Fig. 4 Structure of the DLNN algorithm

$$NN_i = \tau_i e^{\left(\frac{\|cp_c - cp_p\|^2}{2\sigma_i^2}\right)} \tag{28}$$

The learning rate is exhibited as τ_i , the coordinate positions of the current neuron and the previous neuron are depicted as cp_c and cp_p , and the neighbourhood radius's width is proffered as σ_i . Consequently, the HL output is inputted to the OL, in which by pondering the kernelized function, the AF is obtained as:

$$ol_i = ib_s + \sum_{i=1}^n \gamma_i (hd_i \epsilon_i) \tag{29}$$

Herein, the output unit is specified as ol_i and the kernelized function is signified as γ_i . Finally, the loss function is measured as:

$$ls = (ta - ol_i) \tag{30}$$

The loss function is exhibited as ls , the output unit is illustrated as ol_i and the network's target output is depicted as ta . The obtained loss is verified whether it is matched with the specified threshold value or not. If it is matched, then the output is mentioned as the final output or else the weight value is altered again.

4. Result and discussion

Here, the proposed research technique's performance is examined. MATLAB has been employed to execute the IoT platform design and SPVPP DA. The proposed model utilizes publicly obtainable data for the performance investigation.

4.1 Performance analysis of displayed values

The sensed values as of the sensor are examined here. Regarding temperature, time, GSR, along with WS, the SPG is evaluated.

Regarding diverse time instants, Fig. 5 exhibits the SPG. From 6:00 to 19:00, it can perceive. For SPP monitoring, it is one of the other vital parameters. The solar generation level is higher at 12:00 and 13:00. The SPG is lesser for the remaining time. Regarding the temperature variation, Fig. 6 presents the SPG. It is noticed that as the temperature increases, the SPG also increases. The generated solar power is 3500 kW if the temperature is 270 °C. The solar power is 1725 kW if the temperature is 25.50 °C. Therefore, it exhibits that the temperature change offers a large effect on SPG.

Regarding WS changes, Fig. 7 displays the SPG. For power plant monitoring, it is a significant parameter. For the WS's variation, the vast difference is not exhibited. SPG is varied slightly. However, the higher SPG is acquired for the WS of 16.3 km/h. Fig. 8 exhibits the SPG regarding GSR. It is observed that the SPG is increased as the average GSR is increased. The generated solar power is 3900 kW when the average GSR is 5. Conversely, the solar power generated is 995 kW when the average GSR is 1.5. Hence, on SPG, there is a larger effect of average GSR.

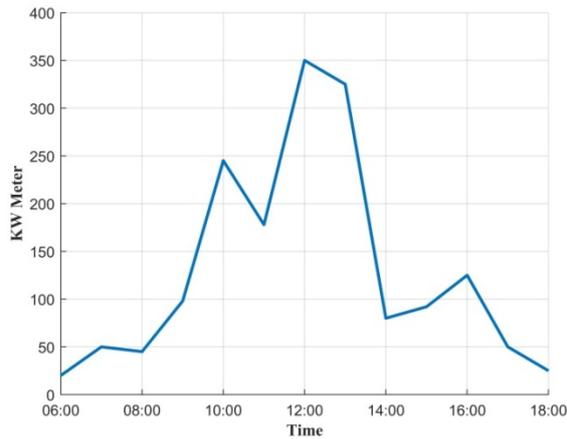


Fig. 5 Analysis of SPG at different time instants

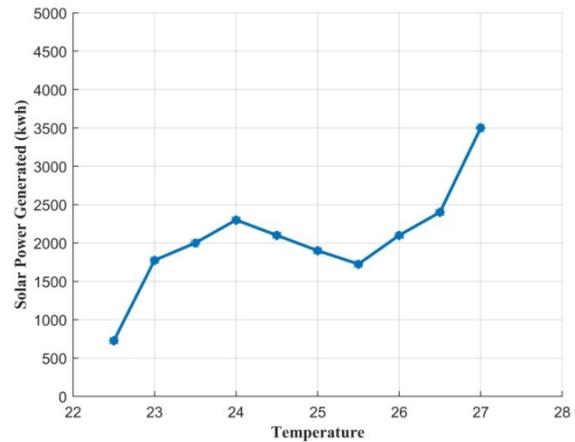


Fig. 6 Analysis of SPG with respect to temperature

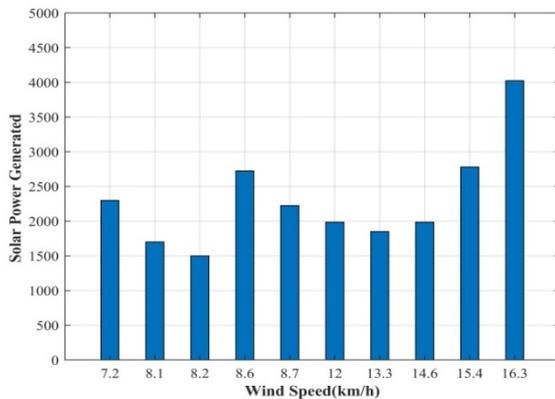


Fig. 7 Analysis of SPG with respect to WS

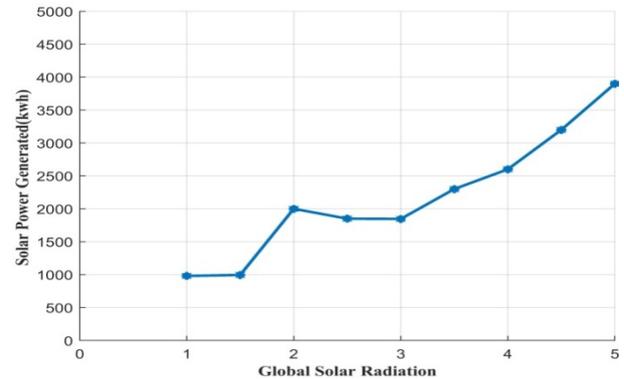


Fig. 8 Power generation analysis with respect to the global solar radiation changes

4.2 Performance analysis of solar power plant monitoring

Here, regarding precision, accuracy, F-measure, and recall metrics, the performance of the SPP monitoring by the GKDLNN is evaluated with the prevailing Support Vector Machine (SVM), ANN, Convolutional NN (CNN), and DLNN algorithms. Furthermore, regarding fitness vs. iteration analysis, the TSSO centered FS process’s performance is analogized with prevailing GA, Particle Swarm Optimization (PSO), Gray Wolf Optimization (GWO), and SO.

Regarding precision, accuracy, F-measure, and recall metrics, the GKDLNN centered SPP monitoring along with the prevailing algorithm-centered monitoring system’s performance is exhibited in Table 1. When analogized to the other prevailing algorithms and the GKDLNN algorithm, the prevailing SVM acquires extremely poor performance in this investigation. On considering the prevailing method, the GKDLNN algorithm acquires higher-level performance. Here, the prevailing DLNN surpassed the other prevailing algorithms; however, it is also lesser than the GKDLNN algorithm.

Table 1 Performance analysis of proposed GKDLNN with the existing algorithm based on performance metrics

Performance Metrics	Proposed GKDLNN	DLNN	CNN	ANN	SVM
Accuracy	96.5	91	88	83.2	79
Precision	94	89.3	85	81.3	78
Recall	94.5	90.3	86	81.5	79.2
F-Measure	94.24	89.79	85.49	81.39	78.59

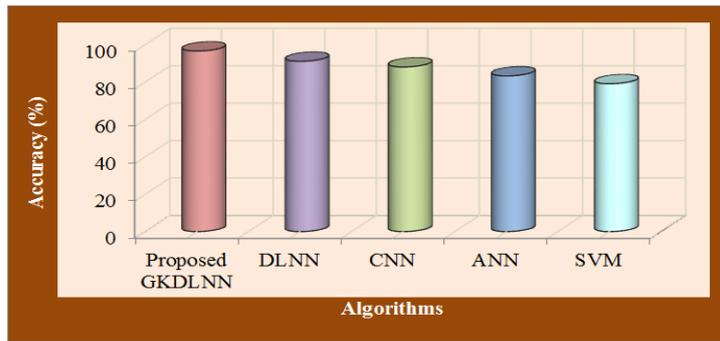
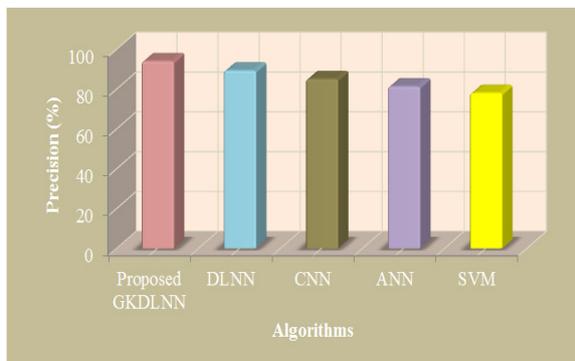
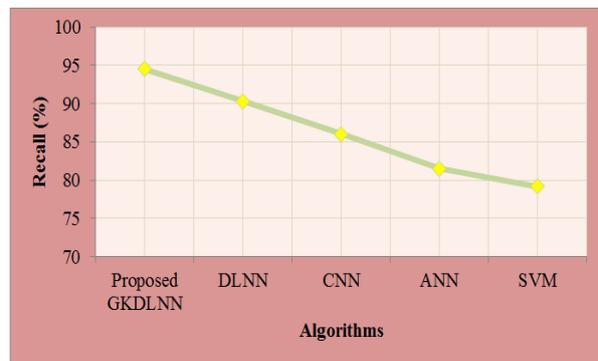


Fig. 9 Graphical representation of accuracy analysis for the proposed and existing algorithms

In Fig. 9, the pictorial depiction of accuracy investigation for the GKDLNN centered SPP monitoring is analogized with the prevailing algorithm-centered monitoring like SVM, ANN, CNN, and DLNN algorithms. The common assessment metric for the predictions is the accuracy metric. It is the number of correct predictions made as a ratio of every prediction done. Accuracy of 96.5 % is acquired by the GKDLNN methodology, which is higher than the prevailing study techniques. For DLNN, CNN, ANN, and SVM, the prevailing technique’s accuracy values are 91%, 88%, 83.2%, and 79 %, respectively.



(a)



(b)

Fig. 10 Demonstrate the performance of the proposed and existing algorithms in terms of (a) precision and (b) recall metrics

Regarding (a) precision and (b) recall metrics, the GKDLNN method's performance investigation with the prevailing algorithms is displayed in Fig. 10. The two very vital model evaluation metrics are recall and precision. Precision implies the percentage of the outcomes which is pertinent, recall implies the percentage of total pertinent outcomes precisely classified by the algorithm. The GKDLNN algorithm accomplished 94 % of precision and 94.5 % of recall; however, the prevailing algorithms like DLNN, CNN, ANN, and SVM have 89.3% and 90.3%, 85 % and 86 %, 81.3 %, and 81.5 %, and 78 % and 79.2 %, respectively. When analogized to the prevailing study techniques, the GKDLNN accomplishes higher performance in this analysis.

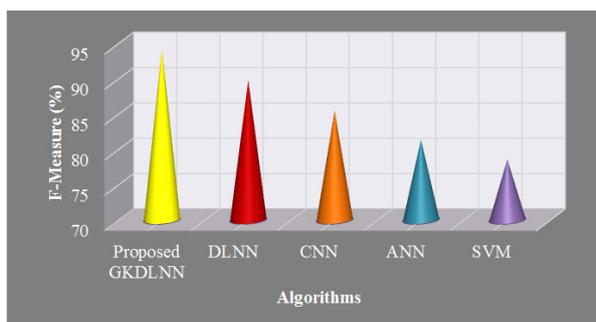


Fig. 11 F-Measure analysis

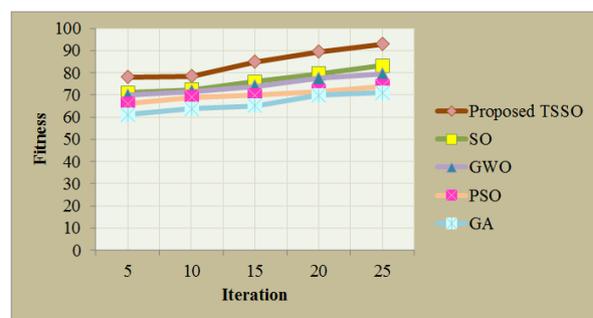


Fig. 12 Fitness vs. iteration analysis

The GKDLNN and prevailing technique's F-measure are exhibited in Fig. 11. F-Measure is the harmonic mean value of the precision and the recall metric. A higher value is also accomplished by the F-measure. The GKDLNN method accomplishes better performance centered on the precision together with recall metrics, so the F-measure-centered analysis is better for the GKDLNN method. Here, an F-measure of 94.24 % is acquired by the GKDLNN technique. The prevailing algorithms' F-measures are 89.79 % for DLNN, 85.49 % for CNN, 81.39 % for ANN, and 78.59 % for SVM. Therefore, the GKDLNN technique's performance surpassed the prevailing techniques as per the analysis.

The TSSO algorithm's fitness vs. iteration investigation with the prevailing GWO, GA, SO, and PSO algorithms are displayed in Fig. 12. The TSSO algorithm's FV is 85 when the iteration count is 15; however, the prevailing techniques namely SO, GWO, PSO, and GA have 76, 73.89, 69.9, and 65.1, respectively. When analogized to the prevailing technique, the TSSO has a greater fitness value for the other iterations also.

5. Conclusion

In this work DA of SPVPPs, an IoT platform's design and development are presented. IoT design and DA are the two phases of this study method. The LPWAN communication layer is employed with Sigfox communication technology, and diverse sensors are employed in the IoT designing phase. The PE, EL, and the fault are evaluated in the DA phase. The GKDLNN algorithm is primarily utilized for analytics. The publicly obtainable dataset is wielded for the performance study. In the outcome analysis phase, the sensed values as of the sensor are plotted, and regarding the recall, precision, F-measure, and accuracy, the GKDLNN algorithm's performance is verified with the prevailing DNN, ANN, DLNN, and SVM models. The GKDLNN methodology acquires the greatest accuracy of 96.5 %. Furthermore, regarding fitness, the TSSO's performance is scrutinized with the prevailing GA, PSO, GWO, and SO. Better performance is achieved by the TSSO algorithm. Therefore, for the SPVPP monitoring, the recommended technique-centered monitoring and the chosen devices are highly cooperative. For ameliorating the system's performance, the proposed mechanism can be expanded encompassing more sensors, the latest devices, along with advanced algorithms.

References

- [1] Shakya, S. (2021). A self monitoring and analyzing system for solar power station using IoT and data mining algorithms, *Journal of Soft Computing Paradigm*, Vol. 3, No. 2, 96-109, doi: [10.36548/jscp.2021.2.004](https://doi.org/10.36548/jscp.2021.2.004).
- [2] Priharti, W., Rosmawati, A.F.K., Wibawa, I.P.D. (2019). IoT based photovoltaic monitoring system application, *Journal of Physics: Conference Series*, Vol. 1367, Article No. 012069, doi: [10.1088/1742-6596/1367/1/012069](https://doi.org/10.1088/1742-6596/1367/1/012069).
- [3] Kekre, A., Gawre, S.K. (2017). Solar photovoltaic remote monitoring system using IOT, In: *Proceedings of 2017 International Conference on Recent Innovations in Signal processing and Embedded Systems (RISE)*, Bhopal, India, 619-623, doi: [10.1109/RISE.2017.8378227](https://doi.org/10.1109/RISE.2017.8378227).
- [4] Sathesh, M., Mohana, J. (2020). IoT based monitoring and load sharing of grid connected solar PV system, *Journal of Critical Reviews*, Vol. 7, No. 6, 2223-2232, doi: [10.31838/jcr.07.6.346](https://doi.org/10.31838/jcr.07.6.346).
- [5] Madhubala, S., Nachammai, R., Nandhini, I., Preethisha, A.M., Janet Paulin, J. (2018). Solar power based remote monitoring and control of industrial parameters using IoT, *International Research Journal of Engineering and Technology*, Vol. 5, No. 3, 3231-3236.
- [6] Kavitha, V., Malathi, V. (2019). A smart solar PV monitoring system using IOT, In: *Proceedings of SRAIC 2019: First International Conference on Secure Reconfigurable Architectures and Intelligent Computing*, Tiruchirappalli Trichy, India, 1-33, doi: [10.5121/csit.2019.91502](https://doi.org/10.5121/csit.2019.91502).
- [7] Ansari, S., Ayob, A., Hossain Lipu, M.S., Saad, M.H.M., Hussain, A. (2021). A review of monitoring technologies for solar PV systems using data processing modules and transmission protocols: Progress, challenges and prospects, *Sustainability*, Vol. 13, No. 15, Article No. 8120, doi: [10.3390/su13158120](https://doi.org/10.3390/su13158120).
- [8] Babu, V., Ahmed, K.S., Shuaib, Y.M., Mani, M. (2021). A novel intrinsic space vector transformation based solar fed dynamic voltage restorer for power quality improvement in distribution system, *Journal of Ambient Intelligence and Humanized Computing*, Vol. 12, No. 1, 897-919, doi: [10.1007/s12652-020-02831-0](https://doi.org/10.1007/s12652-020-02831-0).
- [9] Pereira, R.I.S., Jucá, S.C.S., Carvalho, P.C.M. (2019). IoT embedded systems network and sensors signal conditioning applied to decentralized photovoltaic plants, *Measurement*, Vol. 142, 195-212, doi: [10.1016/j.measurement.2019.04.085](https://doi.org/10.1016/j.measurement.2019.04.085).

- [10] Adhya, S., Saha, D., Das, A., Jana, J., Saha, H. (2016). An IoT based smart solar photovoltaic remote monitoring and control unit, In: *Proceedings of 2nd International Conference on Control, Instrumentation, Energy and Communication (CIEC)*, Kolkata, India, 432-436, doi: [10.1109/CIEC.2016.7513793](https://doi.org/10.1109/CIEC.2016.7513793).
- [11] Badave, P.M., Karthikeyan, B., Badave, S.M., Mahajan, S.B., Sanjeevikumar, P., Gill, G.S. (2018). Health monitoring system of solar photovoltaic panel: An Internet of Things application, In: SenGupta, S., Zobaa, A., Sherpa, K., Bhoi, A. (eds.), *Advances in smart grid and renewable energy. Lecture notes in electrical engineering, Vol 435*, Springer, Singapore, 347-355, doi: [10.1007/978-981-10-4286-7_34](https://doi.org/10.1007/978-981-10-4286-7_34).
- [12] Shrihariprasath, B., Rathinasabapathy, V. (2016). A smart IoT system for monitoring solar PV power conditioning unit, In: *Proceedings of 2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave)*, Coimbatore, India, 1-5, doi: [10.1109/STARTUP.2016.7583930](https://doi.org/10.1109/STARTUP.2016.7583930).
- [13] Ammar, M., Russello, G., Crispo, B. (2018). Internet of Things: A survey on the security of IoT frameworks, *Journal of Information Security and Applications*, Vol. 38, 8-27, doi: [10.1016/j.jisa.2017.11.002](https://doi.org/10.1016/j.jisa.2017.11.002).
- [14] Karbhari, G.V., Nema, P. (2020). IoT & machine learning paradigm for next generation solar power plant monitoring system, *International Journal of Advanced Science and Technology*, Vol. 29, No. 3, 7287-7295.
- [15] Babu, V., Basha, S.S., Shuaib, Y.M., Manikandan, M., Enayathali, S.S. (2019). A novel integration of solar fed dynamic voltage restorer for compensating sag and swell voltage in distribution system using enhanced space vector pulse width modulation (ESVPWM), *Universal Journal of Electrical and Electronic Engineering*, Vol. 6, No. 5, 329-350, doi: [10.13189/ujeee.2019.060504](https://doi.org/10.13189/ujeee.2019.060504).
- [16] Manikandan, M., Basha, A.M. (2016). ODFP: Optimized dual fuzzy flow controller based voltage sag compensation for SMES-based DVR in power quality applications, *Circuits and Systems*, Vol. 7, No. 10, 2959-2974, doi: [10.4236/cs.2016.710254](https://doi.org/10.4236/cs.2016.710254).
- [17] Praveen Kumar, T., Ganapathy, S., Manikandan, M. (2022). Improvement of voltage stability for grid connected solar photovoltaic systems using static synchronous compensator with recurrent neural network, *Electrical Engineering & Electromechanics*, No. 2, 69-77, doi: [10.20998/2074-272X.2022.2.10](https://doi.org/10.20998/2074-272X.2022.2.10).
- [18] Ansari, S., Ayob, A., Hossain Lipu, M.S., Saad, M.H.M., Hussain, A. (2020). Comparison of the IoT based modules for solar PV environment: A review, In: *Proceedings of 2020 IEEE Student Conference on Research and Development (SCoReD)*, Batu Pahat, Malaysia, 401-405, doi: [10.1109/SCoReD50371.2020.9250946](https://doi.org/10.1109/SCoReD50371.2020.9250946).
- [19] Tellawar, M., Chamat, N. (2019). An exclusive review on IoT based solar photovoltaic remote monitoring and controlling unit, *International Research Journal of Engineering and Technology (IRJET)*, Vol. 6, No. 5, 1520-1525.
- [20] Karbhari, G.V., Nema, P. (2019). Digital control system for solar power plant using IoT, *International Journal of Recent Technology and Engineering (IJRTE)*, Vol. 8, No. 2, 3394-3396, doi: [10.35940/ijrte.A1181.078219](https://doi.org/10.35940/ijrte.A1181.078219).
- [21] Sutikno, T., Purnama, H.S., Pamungkas, A., Fadlil, A., Alsofyani, I.M., Jopri, M.H. (2021). Internet of things-based photovoltaics parameter monitoring system using NodeMCU ESP8266, *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 11, No. 6, 5578-5587, doi: [10.11591/ijece.v11i6.pp5578-5587](https://doi.org/10.11591/ijece.v11i6.pp5578-5587).
- [22] Sarswat, S., Yadav, I., Maurya, S.K. (2019). Real time monitoring of solar PV parameter using IoT, *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, Vol. 9, No. 1S, 267-271, doi: [10.35940/ijitee.A1054.1191S19](https://doi.org/10.35940/ijitee.A1054.1191S19).
- [23] Gupta, V., Sharma, M., Pachauri, R.K., Babu, K.N.D. (2020). A low-cost real-time IOT enabled data acquisition system for monitoring of PV system, *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, Vol. 43, No. 20, 2529-2543, doi: [10.1080/15567036.2020.1844351](https://doi.org/10.1080/15567036.2020.1844351).
- [24] Sathish, C., Chidambaram, I.A., Manikandan, M. (2022). Reactive power compensation in a hybrid renewable energy system through fuzzy based boost converter, *Problemele Energeticii Regionale*, No. 53, 10-26, doi: [10.52254/1857-0070.2022.1-53.02](https://doi.org/10.52254/1857-0070.2022.1-53.02).
- [25] Kandimalla, J., Kishore, D.R. (2017). Web based monitoring of solar power plant using open source IOT platform Thingspeak and Arduino, *International Journal for Modern Trends in Science and Technology*, Vol. 3, No. 4, 16-21.
- [26] Almonacid-Olleros, G., Almonacid, G., Fernandez-Carrasco, J.I., Espinilla-Estevez, M., Medina-Quero, J. (2020). A new architecture based on IoT and machine learning paradigms in photovoltaic systems to nowcast output energy, *Sensors*, Vol. 20, No. 15, Article No. 4224, doi: [10.3390/s20154224](https://doi.org/10.3390/s20154224).
- [27] Ramamurthi, P.V., Nadar, E.R.S. (2022). IoT-based energy monitoring and controlling of an optimum inclination angle of the solar panels, *IETE Journal of Research*, Vol. 68, No. 4, 3108-3118, doi: [10.1080/03772063.2020.1754301](https://doi.org/10.1080/03772063.2020.1754301).
- [28] Paredes-Parra, J.M., García-Sánchez, A.J., Mateo-Aroca, A., Molina-García, Á. (2019). An alternative Internet-of-Things solution based on LoRa for PV power plants data monitoring and management, *Energies*, Vol. 12, No. 5, Article No. 881, doi: [10.3390/en12050881](https://doi.org/10.3390/en12050881).
- [29] Cheddadi, Y., Cheddadi, H., Cheddadi, F., Errahimi, F., Es-sbai, N. (2020). Design and implementation of an intelligent low-cost IoT solution for energy monitoring of photovoltaic stations, *SN Applied Sciences*, Vol. 2, No. 7, Article No. 1165, doi: [10.1007/s42452-020-2997-4](https://doi.org/10.1007/s42452-020-2997-4).
- [30] Samkria, R., Abd-Elnaby, M., Singh, R., Gehlot, A., Rashid, M., Aly, M.H., El-Shafai, W. (2021). Automatic PV grid fault detection system with IoT and LabVIEW as data logger, *Computers, Materials and Continua*, Vol. 69, No. 2, 1709-1723, doi: [10.32604/cmc.2021.018525](https://doi.org/10.32604/cmc.2021.018525).
- [31] Oukennou, A., Berrar, A., Belbhar, I., El Hamri, N. (2019). Low cost IoT system for solar panel power monitoring, In: *Colloque sur les Objets et systèmes Connectés, Ecole Supérieure de Technologie de Casablanca, Institut Universitaire de Technologie d'Aix-Marseille*, Casablanca, Morocco.
- [32] Shapsough, S., Takroui, M., Dhaouadi, R., Zualkernan, I.A. (2021). Using IoT and smart monitoring devices to optimize the efficiency of large-scale distributed solar farms, *Wireless Networks*, Vol. 27, 4313-4329, doi: [10.1007/s11276-018-01918-z](https://doi.org/10.1007/s11276-018-01918-z).

- [33] Babu, V., Ahmed, K.S., Shuaib, Y.M., Manikandan, M. (2021). Power quality enhancement using dynamic voltage restorer (DVR)-based predictive space vector transformation (PSVT) with proportional resonant (PR)-controller, *IEEE Access*, Vol. 9, 155380-155392, doi: [10.1109/ACCESS.2021.3129096](https://doi.org/10.1109/ACCESS.2021.3129096).
- [34] Pulungan, A.B., Risfendra, R., Purwanto, W., Maksum, H., Setiawan, O. (2020). Design and development of real time monitoring single axis solar tracker by using internet of things, *International Journal of Geomate*, Vol. 18, No. 69, 81-87, doi: [10.21660/2020.69.25863](https://doi.org/10.21660/2020.69.25863).
- [35] Reddy, S.G., Ganapathy, S., Manikandan, M. (2022). Three phase four switch inverter based DVR for power quality improvement with optimized CSA approach, *IEEE Access*, Vol. 10, 72263-72278, doi: [10.1109/ACCESS.2022.3188629](https://doi.org/10.1109/ACCESS.2022.3188629).
- [36] Reddy, S.G., Ganapathy, S., Manikandan, M. (2022). Power quality improvement in distribution system based on dynamic voltage restorer using PI tuned fuzzy logic controller, *Electrical Engineering & Electromechanics*, No. 1, 44-50, doi: [10.20998/2074-272X.2022.1.06](https://doi.org/10.20998/2074-272X.2022.1.06).

Comparing Fault Tree Analysis methods combined with Generalized Grey Relation Analysis: A new approach and case study in the automotive industry

Shi, J.L.^a, Lu, Z.C.^{a,*}, Xu, H.H.^a, Ren, M.M.^a, Shu, F.L.^a

^aSchool of Mechanical Engineering and Automation, Dalian Polytechnic University, Dalian, P.R. China

ABSTRACT

The failure modes of products gradually show a diversified trend with the precision and complexity of the product structure. The combination of fault tree analysis and generalized grey relational analysis is widely used in the fault diagnosis of complex systems. In this study, we utilize a method that combines fault tree analysis and generalized grey relational analysis. This method is applied to diagnose the *Expansion Adhesive Debonding* fault of automobile doors. Then, we analyse and compare the differences in actual fault diagnosis results. The comparison involves three analysis methods: Fault Tree Analysis combined with Absolute Grey Relation Analysis (F-AGRA), Fault Tree Analysis combined with Relative Grey Relation Analysis (F-RGRA), and Fault Tree Analysis combined with Comprehensive Grey Relation Analysis (F-CGRA). Subsequently, we compare the findings with actual production results. This comparison allows us to discuss the differences between the three methods in the fault diagnosis of complex systems. We also discuss the application occasions of these methods. This study will provide a new method for fault analysis and fault diagnosis in the actual production of the automobile manufacturing industry. This method can eliminate faults effectively and accurately and improve product quality and productivity.

ARTICLE INFO

Keywords:

Fault tree analysis (FTA);
Generalized Grey Relation Analysis (GGRA);
Failure mode;
Fault diagnosis;
Complex system;
Fault Tree Analysis combined with Absolute Grey Relation Analysis (F-AGRA);
Fault Tree Analysis combined with Relative Grey Relation Analysis (F-RGRA);
Fault Tree Analysis combined with Comprehensive Grey Relation Analysis (C-GRA)

*Corresponding author:

577345879@qq.com
(Lu, Z.C.)

Article history:

Received 5 November 2023
Revised 19 December 2023
Accepted 21 December 2023



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

1. Introduction

The diversified needs of customers and the increasingly improved automation and intelligent production have brought about product function diversity and fast-paced production. However, the diagnosis and analysis of the complicated and diversified failure modes have led to new challenges. The products in the automobile manufacturing industry exhibit more complex failure modes, more diverse fault causes, and a stronger correlation between fault causes than those in other industries. The simple fault analysis method cannot solve the complex fault problem.

Therefore, a fault analysis method for complex system analysis that can accurately, quickly, and efficiently determine the cause of product failure and take corresponding troubleshooting measures is urgently needed.

Complex systems are generally composed of various interacting components with complex relationships. This scenario also leads to system failures that typically involve multiple levels of organization and interrelated factors. As the system's complexity increases, the interactions between components and the complexity of the hierarchical structure and interrelationships between components increase. Thus, the complexity of analysis increases with the increase in system complexity.

Fault Tree Analysis (FTA) is currently recognized as an effective technique for determining and analysing failure modes of complex products or systems [1]. This kind of inverted-tree logical cause-and-effect diagram is established for studying the system's functional failure [2]. This analysis method has been widely used in the mechanical engineering [3], medicine [4], chemistry [5], transportation [6], construction [7], aviation [8] and others for qualitative and quantitative fault diagnoses. It plays an important role in system reliability analysis, safety analysis, and risk assessment [9]. However, the correlation between the basic and top events is ignored in the practical application of FTA. The components' potential and relative failure events in complex systems often exist simultaneously. Thus, the system failure and the resulting failure are usually uncertain [10]. Moreover, the correlation information between them is lacking. Thus, inaccurate results easily occur when FTA is used to analyse complex systems. Therefore, the correlation between fault events and system failures must be analysed by other effective methods combined with FTA. Grey relational analysis (GRA) is one of the most widely used methods.

Based on the sample data of various factors, GRA uses grey relational degrees to describe the strength, size, and order of the relationships between factors [11-12]. The basic idea is that if the changing trends of the two factors are consistent, they can be considered relatively related. Otherwise, the correlation between them is small. Therefore, according to the basic idea of GRA, it can be applied to comprehensive evaluation. The evaluated object and the ideal object are regarded as two systems, and the advantages and disadvantages of the evaluated object can be obtained by calculating the correlation degree of the two systems. The greater the correlation degree of the two systems is, the better the evaluated object is. Alternatively, the evaluated object and the negative ideal object can be regarded as two systems. The greater the correlation between the two systems is, the more similar the failure mode of the evaluated object is to the negative ideal solution [13].

The combination of FTA and GRA has been extensively researched in recent years. Pang calculated the relationship between basic event and top event failure rate trend in the electromagnetic machining process by combining T-S FTA and GRA. In this way, they realized the targeted repair of the fault system in the electromagnetic machining process and achieved the purpose of quality control [14]. Tien optimized the process parameters by combining fuzzy GRA and particle swarm optimization based on the Taguchi experiment, considerably improving the performance indicators of materials [15]. Applying this machine learning method to fault diagnosis enables effective handling of large-scale, high-dimensional data and the discovery of fault features and trends hidden behind extensive datasets. This method is crucial for fault diagnosis in complex systems. It can analyse and learn from data in real time. Thus, it has the potential for real-time fault diagnosis. It can also provide objective, data-driven analyses, thereby reducing subjective interventions. Wang *et al.* introduced grey theory and fuzzy theory into the FTA to analyse the wind turbine. They proposed the gearbox transmission system based on grey theory and T-S fuzzy FTA to solve uncertain failure probability [16]. Wang introduced the method of combining GRA and FTA in the analysis and diagnosis of the loosening accident in a mine hoist. He analysed the possibility of various fault modes in the FTA of excessive looseness in a mine hoist [17]. Zhang *et al.* proposed a failure mode analysis method to improve the generalized grey relational fault tree for mine car fall accidents. They used this method to determine and rank the possibility of occurrence of various mine car fall failure modes [18]. Chen *et al.* proposed the LSSC reliability diagnosis method for hazardous materials by improving the generalized grey relation fault tree. This method can quickly identify the accident factors of the key chemical logistics operation system [19].

The works above show that FTA is mainly used for deterministic fault analysis, whereas GRA is primarily used for uncertain fault analysis. Deterministic methods have low data requirements and may be fast because they focus on deterministic faults. On the contrary, uncertain fault analysis methods are comprehensive and can handle the interrelationships among multiple faults in a system. They usually have high data requirements to capture the relationships and uncertainties. This comprehensiveness can enhance problem identification accuracy but may require increased computational resources and time.

Therefore, combining FTA and GRA methods can leverage the advantages of deterministic and nondeterministic methods for fault analysis and diagnosis. This approach allows for the rapid and effective analysis of the interrelationships among multiple fault causes in system failures. Therefore, applying the combination of FTA and GRA to solve product failures encountered in industrial product manufacturing has a great generalization value. Chen *et al.* determined the weak link in the LSSC system for a hazardous chemical by combining generalized GRA with a fault tree [19]. This method can be extended to product failure analysis in industrial product manufacturing.

At present, the analysis methods combining generalized GRA with fault trees commonly include FTA combined with Absolute Grey Relation Analysis (called F-AGRA), FTA combined with Relative Grey Relation Analysis (called F-RGRA), and FTA combined with Comprehensive Grey Relation Analysis (called F-CGRA). These three methods have their respective advantages, disadvantages, and applications. However, the current study has no special research and focused discussion on the applicability and accuracy of the three methods in the failure mode of complex products. As a result, the failure analysis using the combination of generalized GRA and FTA cannot quickly and accurately choose among the three methods. Thus, the production efficiency is affected.

The present study established a fault tree by taking the typical automobile doors' *Expansion Adhesive Debonding* fault encountered in the production of G Group as the research object to make up for the shortcomings of the studies above. The automobile doors' *Expansion Adhesive Debonding* fault was also analysed based on F-AGRA, F-RGRA, and F-CGRA methods. We followed the actual problem-solving process in the production site to understand the causes and final solutions of the failure in the actual production. In this study, the theoretical analysis of the three methods was compared with that of the troubleshooting methods in the actual production. We studied and analysed the reasons for the differences between the three methods to choose the suitable fault analysis method. This study can provide an efficient and accurate idea of analysing the causes of faults for enterprise fault improvement to solve system faults quickly.

2. Expansion Adhesive Debonding fault analysis

In the disassembly and self-inspection processes for the exit door assembly of a certain model of G Group, the expansion adhesive filled between the inner and outer plates of the automobile doors did not fit the surface of the sheet metal parts. Moreover, the expansion adhesive section was neat. No attachment was found between the door plate and the door plate, and a large gap exists. According to the "Specification for Complete Destruction of Body in White Connection Quality" of the group, the expansion adhesive should have adhesion traces on all the panels on both sides of the split. Otherwise, the expansion adhesive may have insufficient filling, leading to noise caused by the vibration during driving. The noise becomes increasingly obvious with the increase in speed, seriously affecting user comfort. Therefore, the *Expansion Adhesive Debonding* fault of the automobile door is defined as a system fault, and the F-AGRA, F-RGRA, and F-CGRA methods are adopted for analysis. The results are compared with the actual practice to analyse the reasons for the differences and applicable occasions of the three methods.

2.1 Fault analysis process based on Fault Tree Analysis and Generalized Grey Relation Analysis

The Fault Tree and Grey Relation Analysis are used to analyse the problem. The specific process is as Fig. 1.

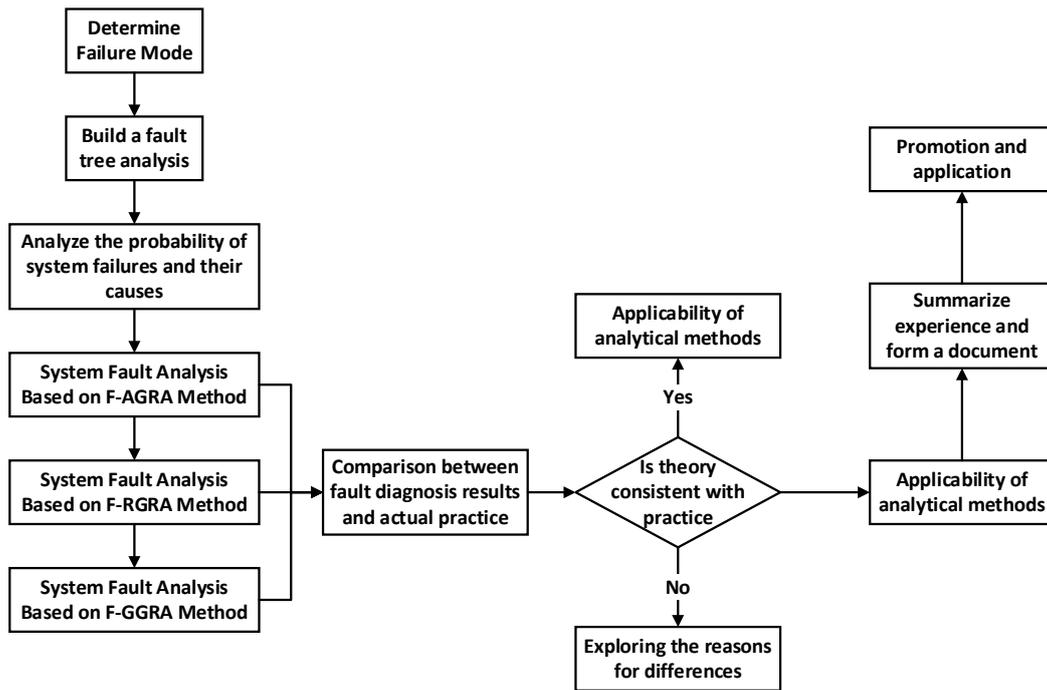


Fig. 1 Fault analysis process based on Fault Tree Analysis and Generalized Grey Relation Analysis

2.2 Constructing the Fault Tree Analysis construction

FreeFta software is used to construct the fault tree of automobile doors' *Expansion Adhesive Debonding*, as shown in Fig. 2. T represents the fault tree top event, M_i represents the intermediate event, and N_i represents the basic event. In this case, the *Expansion Adhesive Debonding* fault of the automobile door is defined as the top event, which is a system fault. The settings of top events, intermediate events, and basic events are shown in Table 1.

2.3 Failure probability assessment

G Group established a fault analysis team comprising field technicians and engineers to solve the *Expansion Adhesive Debonding* fault of automobile doors by evaluating the possible risk probability of the basic events of the fault tree. Moreover, Table 2 is the standard table of the problem level and probability assessment found in the production process investigation, and Table 3 is the probability assessment table of the basic event N_1 - N_{15} of the *Expansion Adhesive Debonding* fault of automobile doors.

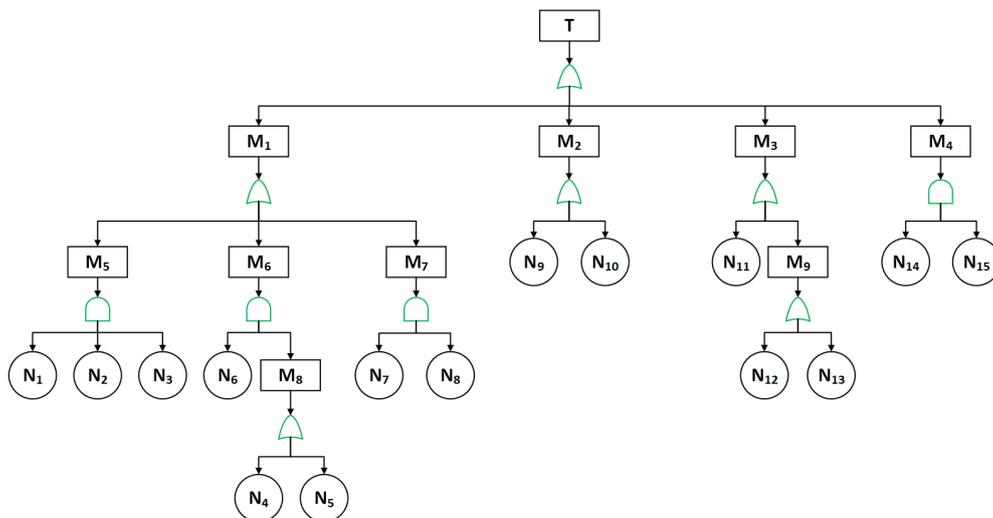


Fig. 2 Fault Tree Analysis of *Expansion Adhesive Debonding* of automobile doors

Table 1 Events represented by letters in Fault Tree Analysis

Event Type	Event Code	Fault event
Top event	<i>T</i>	Automobile doors <i>Expansion Adhesive Debonding</i>
Intermediate events	<i>M</i> ₁	Excessive spacing between sheet metal parts
	<i>M</i> ₂	Gluing equipment error
	<i>M</i> ₃	The properties of the expansion adhesive do not meet the requirements
	<i>M</i> ₄	Improper transportation and storage methods for assembly parts
	<i>M</i> ₅	Part size deviation exceeding standard
	<i>M</i> ₆	Excessive deformation of parts
	<i>M</i> ₇	Part position deviation exceeds the standard
	<i>M</i> ₈	Part deformation caused by the process of transferring parts
	<i>M</i> ₉	The properties of the expansion adhesive are not sufficient to meet the needs of sea transportation
Elementary event	<i>N</i> ₁	Dimensional deviation of outer door panel
	<i>N</i> ₂	Dimensional deviation of door inner panel
	<i>N</i> ₃	Dimensional deviation of reinforced beams
	<i>N</i> ₄	Deformation of the reinforced beam caused by taking it
	<i>N</i> ₅	Deformation of parts caused by handling after welding completion
	<i>N</i> ₆	Overpressure of fixture pressure arm
	<i>N</i> ₇	Poor consistency of solder joints leads to fluctuations in part position
	<i>N</i> ₈	Excessive wear of fixture positioning pins
	<i>N</i> ₉	Deviation of gluing point position
	<i>N</i> ₁₀	Insufficient glue application
	<i>N</i> ₁₁	Insufficient viscosity of expansion adhesive
	<i>N</i> ₁₂	Deterioration of expanded adhesive due to prolonged exposure time
	<i>N</i> ₁₃	Expansion adhesive deteriorates under high temperature and humidity conditions
	<i>N</i> ₁₄	The coated expansion adhesive is shipped by sea to M country for baking
	<i>N</i> ₁₅	Improper storage method for expansion adhesive during transportation

Table 2 Standard table of problem level and probability assessment found in production process investigation

Level	Probability of failure occurrence	
	Possibility of occurrence	Probability of occurrence value
one	great	0.9
two	more	0.8
three	secondary	0.6
four	less	0.2
five	very small	0.1

Table 3 Probability assessment table of basic event *N*₁-*N*₁₅ of *Expansion Adhesive Debonding* failure of automobile

Elementary event <i>N</i> _{<i>i</i>}	<i>N</i> ₁	<i>N</i> ₂	<i>N</i> ₃	<i>N</i> ₄	<i>N</i> ₅	<i>N</i> ₆	<i>N</i> ₇	<i>N</i> ₈
Probability <i>P</i> (<i>N</i> _{<i>i</i>})	0.05	0.05	0.2	0.05	0.1	0.4	0.3	0.2
Elementary event <i>N</i> _{<i>i</i>}	<i>N</i> ₉	<i>N</i> ₁₀	<i>N</i> ₁₁	<i>N</i> ₁₂	<i>N</i> ₁₃	<i>N</i> ₁₄	<i>N</i> ₁₅	
Probability <i>P</i> (<i>N</i> _{<i>i</i>})	0.2	0.3	0.05	0.7	0.7	0.5	0.4	

3 Analysis based on Fault Tree Analysis of Expansion Adhesive Debonding

3.1 Top event probability

If the basic events in the fault tree are independent of each other, then the probability of the top event can be obtained from Eqs. 1 and 2 [17].

The probability of occurrence of the output event of the and gate structure:

$$P(T) = \prod_{i=1}^n P_i \tag{1}$$

The probability of occurrence of the output event of the or gate structure:

$$P(T) = 1 - \prod_{i=1}^n (1 - P_i) \tag{2}$$

According to Eqs. 1 and 2, the probability of top event occurrence can be calculated as $P(T) = 0.9750$.

3.2 Minimum cut set of fault tree analysis

If a set of basic events occur in the fault tree, the top event is bound to occur. The set of basic events is called the cut set; the minimum cut set represents a minimum failure mode that causes the top event of the fault tree to occur. The study of the minimum cut set helps identify the fault tree’s weaknesses [18].

This study adopts the upward method to calculate the minimum cut sets in the fault tree of the automobile doors *Expansion Adhesive Debonding*. The calculation process is shown in the Eq. 3 [19].

$$\begin{aligned}
 T &= M_1 + M_2 + M_3 + M_4 \\
 &= M_5 + M_6 + M_7 + N_9 + N_{10} + N_{11} + M_9 + N_{14} + N_{15} \\
 &= N_1N_2N_3 + N_6M_8 + N_7N_8 + N_9 + N_{10} + N_{11} + N_{12} + N_{13} + N_{14} + N_{15} \\
 &= N_1N_2N_3 + N_6N_4 + N_6N_5 + N_7N_8 + N_9 + N_{10} + N_{11} + N_{12} + N_{13} + N_{14} + N_{15}
 \end{aligned}
 \tag{3}$$

It should be noted that $N_1N_2N_3$ is a simplified expression of $\{N_1, N_2, N_3\}$, representing a fault event, and the rest are the same. According to Eq. 3, the corresponding minimum cut sets labelled C_1 - C_{11} are: $C_1 = \{N_1, N_2, N_3\}$, $C_2 = \{N_4, N_6\}$, $C_3 = \{N_5, N_6\}$, $C_4 = \{N_7, N_8\}$, $C_5 = \{N_9\}$, $C_6 = \{N_{10}\}$, $C_7 = \{N_{11}\}$, $C_8 = \{N_{12}\}$, $C_9 = \{N_{13}\}$, $C_{10} = \{N_{14}\}$, $C_{11} = \{N_{15}\}$.

3.3 Minimum cut sets probability

Based on the probability of each basic event, each minimum cut set probability is calculated using Eq. 4 [20].

$$P(C_i) = \prod_{j \in C_i} P(x_j)
 \tag{4}$$

According to the calculation of Eq. 4, the minimum cut sets probability of the automobile doors *Expansion Adhesive Debonding* fault is obtained, as shown in Table 4.

Table 4 Minimum cut sets probability of *Expansion Adhesive Debonding* fault in automobile doors

Min. cut set C_i	C_1	C_2	C_3	C_4	C_5	C_6
Probability $P(C_i)$	0.0005	0.02	0.04	0.06	0.2	0.3
Min. cut set C_i	C_7	C_8	C_9	C_{10}	C_{11}	
Probability $P(C_i)$	0.05	0.7	0.7	0.5	0.4	0.05

4. Generalized Grey Relation Analysis

4.1 Importance of basic event

The basic event importance indicates the degree of influence of the basic event on the top event and is calculated by Eq. 5 [19].

$$I_j = \frac{\sum_{j \in C_i} P(C_i)}{P(T)}
 \tag{5}$$

According to the calculation of Eq. 5, the importance of basic events of *Expansion Adhesive Debonding* failure of automobile doors is shown in Table 5.

Table 5 Importance of basic events of *Expansion Adhesive Debonding* failure of automobile doors

Elementary event N_i	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8
Importance $I_j(N_0)$	0.0005	0.0005	0.0005	0.0205	0.0410	0.0615	0.0615	0.0615
Elementary event N_i	N_9	N_{10}	N_{11}	N_{12}	N_{13}	N_{14}	N_{15}	
Importance $I_j(N_0)$	0.2051	0.3077	0.0513	0.7179	0.7179	0.5128	0.4103	

4.2 Fault Analysis based on F-AGRA method

The F-AGRA method can reflect each factor’s degree of influence on the system. The fault tree concept indicates that this degree of influence is the longitudinal difference between the initial feature matrix of the minimum cut sets and the vector of patterns to be examined consisting of the importance of the underlying events. It ultimately reflects the fault patterns’ degree of contribution represented by each minimum cut set to the top event fault.

X_i represents the basic event probability of the i -th minimum cut set, $X = \{X_1, X_2, \dots, X_i, \dots, X_m\}^T$, $X_i = \{x_i(1), x_i(2), \dots, x_i(n)\}$, here m is the number of minimum cut sets, and n is the number of elementary events in the fault system. The characteristic matrix X consisting of subsequences can be expressed as the Eq. 6.

$$X = \begin{Bmatrix} X_1 \\ X_2 \\ \dots \\ X_i \\ \dots \\ X_m \end{Bmatrix} = \begin{Bmatrix} \begin{pmatrix} x_1(1) & x_1(2) & \dots & x_1(n) \\ x_2(1) & x_2(2) & \dots & x_2(n) \\ \dots & \dots & \dots & \dots \\ x_i(1) & x_i(2) & \dots & x_i(n) \\ \dots & \dots & \dots & \dots \\ x_m(1) & x_m(2) & \dots & x_m(n) \end{pmatrix} \end{Bmatrix} \tag{6}$$

In the characteristic matrix X , if a basic event is not in some minimum cut sets, then the vector value for that position is 0. Otherwise, the vector value for that position is the probability of this basic event. Therefore, the characteristic matrix of the basic event probability column can be expressed as the Eq. 7.

$$\begin{pmatrix} 0.05 & 0.05 & 0.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.05 & 0 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.2 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0.2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.05 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4 \end{pmatrix} \tag{7}$$

Define the $X_0 = \{x_0(1), x_0(2), \dots, x_0(n)\}$ as the primary sequence, the vector values corresponding to $x_0(1), x_0(2), \dots, x_0(n)$ are the importance of each basic event respectively, and get pending pattern vector $X_0 = \{0.0005, 0.0005, 0.0005, 0.0205, 0.0410, 0.0615, 0.0615, 0.0615, 0.2051, 0.3077, 0.0513, 0.7179, 0.7179, 0.5128, 0.4103\}$.

Assuming that $X_0 = \{x_0(1), x_0(2), \dots, x_0(n)\}$ is the primary sequence and X_0 and X_i have the same length, the absolute Grey Relation Degree ε_{0i} can be derived from Eq. 8 [21].

$$\left. \begin{aligned} \varepsilon_{0i} &= \frac{1 + |S_0| + |S_i|}{1 + |S_0| + |S_i| + |S_i - S_0|}, \quad i = 1, 2, \dots, n \\ |S_0| &= \left| \sum_{k=1}^{n-1} x_0(k) + \frac{1}{2} x_0(n) \right| \\ |S_i| &= \left| \sum_{k=1}^{n-1} x_i(k) + \frac{1}{2} x_i(n) \right| \\ |S_i - S_0| &= \left| \sum_{k=1}^{n-1} (x_i(k) - x_0(k)) + \frac{1}{2} (x_i(n) - x_0(n)) \right| \end{aligned} \right\} \tag{8}$$

$|S_0|$ is the fluctuation amplitude sum of the primary feature series, $|S_i|$ is the fluctuation amplitude sum of the feature series composed of the factors in the system, and $|S_i - S_0|$ is the difference between the amplitude of fluctuations of factors in the main series and subseries. According to the equation, the absolute grey relational degree of the minimum cut sets for the *Expansion Adhesive Debonding* fault of automobile doors is shown in Table 6.

According to Table 6, we get the Importance Ranking of the minimum cut sets for *Expansion Adhesive Debonding* fault of automobile doors based on F-AGRA method, as:

$$C_8 = C_9 > C_3 > C_4 = C_{10} > C_2 > C_{11} > C_1 = C_6 > C_5 > C_7$$

Fig. 3 is histogram of Absolute Grey Relation of minimum cut sets.

Table 6 Absolute Grey Relation Degree of minimum cut sets for *Expansion Adhesive Debonding* fault of automobile doors

Min cut set C_i	C_1	C_2	C_3	C_4	C_5	C_6
Absolute Grey Relation Degree ε_{0i}	0.6154	0.6371	0.6587	0.6443	0.6010	0.6154
Min cut set C_i	C_7	C_8	C_9	C_{10}	C_{11}	
Absolute Grey Relation Degree ε_{0i}	0.5794	0.6731	0.6731	0.6443	0.6299	

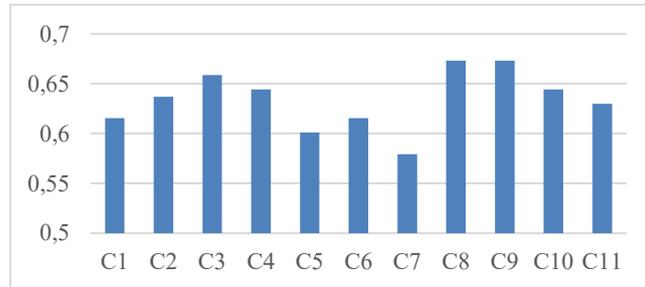


Fig. 3 Histogram of Absolute Grey Relation of minimum cut sets

4.3 Fault Analysis based on F-RGRA method

The F-RGRA method is used to analyse the differences in the longitudinal variation of the subsequence and the primary sequence. Under the fault tree concept, it can be understood as the degree of conformity of the change in each value within each row of the initial feature matrix of the minimum cut sets with the change in the value of the vector of factors of the pattern to be examined consisting of the importance of the basic events. It can eventually reflect the degree of influence of the change in the failure mode of each minimum cut set on the change of the top event.

In the calculation of the relative grey relational degree, the initial value is generally calculated by dividing the first term of the sequence, as shown in the Eq. 9.

$$X'_i = \{x'_i(1), x'_i(2), \dots, x'_i(n)\} = \left\{ \frac{x_i(1)}{x_i(1)}, \frac{x_i(2)}{x_i(1)}, \dots, \frac{x_i(n)}{x_i(1)} \right\} \tag{9}$$

However, this approach becomes problematic if the first term in the data set is 0. Therefore, the equation can express the definition of the initial value by replacing the first term with the largest value in the data set (Eq. 10),

$$X'_i = \{x'_i(1), x'_i(2), \dots, x'_i(n)\} = \left\{ \frac{x_i(1)}{x_i(d)}, \frac{x_i(2)}{x_i(d)}, \dots, \frac{x_i(n)}{x_i(d)} \right\} \tag{10}$$

where $x_i(d) = \max \{x_i\}$. The characteristic matrix of the initial value image can be expressed as Eq. 11.

$$X = \begin{Bmatrix} X'_1 \\ X'_2 \\ \dots \\ X'_i \\ \dots \\ X'_m \end{Bmatrix} = \begin{Bmatrix} \left(\frac{x_1(1)}{x_1(d)} \quad \frac{x_1(2)}{x_1(d)} \quad \dots \quad \frac{x_1(n)}{x_1(d)} \right) \\ \left(\frac{x_2(1)}{x_2(d)} \quad \frac{x_2(2)}{x_2(d)} \quad \dots \quad \frac{x_2(n)}{x_2(d)} \right) \\ \dots \\ \left(\frac{x_i(1)}{x_i(d)} \quad \frac{x_i(2)}{x_i(d)} \quad \dots \quad \frac{x_i(n)}{x_i(d)} \right) \\ \dots \\ \left(\frac{x_m(1)}{x_m(d)} \quad \frac{x_m(2)}{x_m(d)} \quad \dots \quad \frac{x_m(n)}{x_m(d)} \right) \end{Bmatrix} \tag{11}$$

The characteristic matrix of the data after the initial image processing:

$$\begin{pmatrix} 0.25 & 0.25 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.125 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0.67 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{12}$$

Similarly, in the primary sequence $X'_0 = \{x'_0(1), x'_0(2), \dots, x'_0(n)\}$, $x'_0(1), x'_0(2), \dots, x'_0(n)$ can be obtained from Eq. 13.

$$X'_0 = \{x'_0(1), x'_0(2), \dots, x'_0(n)\} = \left\{ \frac{x_0(1)}{x_0(d)}, \frac{x_0(2)}{x_0(d)}, \dots, \frac{x_0(n)}{x_0(d)} \right\} \tag{13}$$

The pending pattern vector $X'_0 = \{0.0007, 0.0007, 0.0007, 0.0256, 0.0571, 0.0857, 0.0857, 0.0857, 0.2857, 0.4286, 0.0715, 1, 1, 0.7143, 0.5715\}$.

If X'_0 and X'_i are the initial value images of X_0 and X_i , then the absolute grey relational degree between X'_0 and X'_i is the relative relation degree between X_0 and X_i , denoted as r_{0i} , which can be obtained from the Eq. 14 [22].

$$\left. \begin{aligned} r_{0i} &= \frac{1 + |S'_0| + |S'_i|}{1 + |S'_0| + |S'_i| + |S'_i - S'_0|}, \quad i = 1, 2, \dots, n \\ |S'_0| &= \left| \sum_{k=1}^{n-1} x'_0(k) + \frac{1}{2} x'_0(n) \right| \\ |S'_i| &= \left| \sum_{k=1}^{n-1} x'_i(k) + \frac{1}{2} x'_i(n) \right| \\ |S'_i - S'_0| &= \left| \sum_{k=1}^{n-1} (x'_i(k) - x'_0(k)) + \frac{1}{2} (x'_i(n) - x'_0(n)) \right| \end{aligned} \right\} \tag{14}$$

$|S'_0|$ is the sum of the rates of change of the volatility of the total characteristic sequence, $|S'_i|$ is the sum of the rates of change of the volatility of the characteristic sequence of the factors in the system, and $|S'_i - S'_0|$ is the sum of the differences in the rates of change in the volatility of the corresponding factors in the primary sequence and the subseries. According to the calculation of the Eq. 14, the relative grey relational degree of the minimum cut sets for the *Expansion Adhesive Debonding* fault of automobile doors is shown in Table 7.

Table 7 Relative Grey Relation Degree of minimum cut sets for *Expansion Adhesive Debonding* faults in automobile doors

Min cut set C_i	C_1	C_2	C_3	C_4	C_5	C_6
Relative Grey Relation Degree r_{0i}	0.7161	0.6756	0.7161	0.7345	0.6621	0.6621
Min cut set C_i	C_7	C_8	C_9	C_{10}	C_{11}	
Relative Grey Relation Degree r_{0i}	0.6621	0.6621	0.6621	0.6621	0.6080	

According to Table 7, we get the Importance Ranking of the minimum cut sets for *Expansion Adhesive Debonding* fault of automobile doors based on F-RGRA method, as:

$$C_4 > C_1 = C_3 > C_2 = C_5 = C_6 = C_7 = C_8 = C_9 = C_{10} > C_{11}$$

Fig. 4 is histogram of Relative Grey Relation of minimum cut sets.

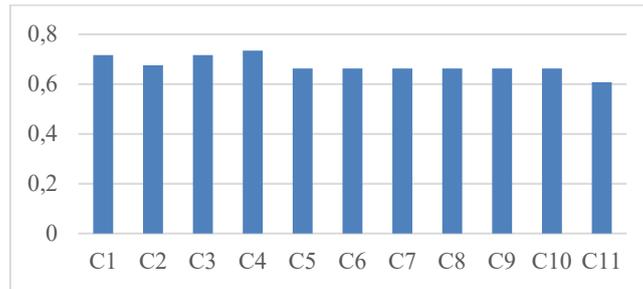


Fig. 4 Histogram of Relative Grey Relation of minimum cut sets

4.4 Fault Analysis based on F-CGRA method

The F-CGRA method is a combination of the F-AGRA method and the F-RGRA method. It integrates the longitudinal differences and rates of change of the selected indicators (called: primary sequence) and related factors (called: subseries). Finally, it realizes the relation analysis of the selected indicators and system-related factors.

In general, the Comprehensive Grey Relation Degree ρ_{0i} is defined by combining the Absolute Grey Relation Degree and the Relative Grey Relation Degree as shown in Eq. 15 [23].

$$\rho_{0i} = \theta_i \varepsilon_{0i} + (1 - \theta_i)r_{0i} \tag{15}$$

In the equation, θ_i is the distribution coefficient, which can be determined by the maximum deviation method proposed by Sun *et al.* as in Eq. 16 [24].

$$\theta_i = \frac{B_{\varepsilon i}}{B_{\varepsilon i} + B_{r i}} \tag{16}$$

Total deviation of Absolute Grey Relation Degree [25]:

$$B_{\varepsilon i} = \sum_{k=1}^n |\varepsilon_{0i} - \varepsilon_{0k}| \tag{17}$$

Total deviation of Relative Grey Relation Degree [25]:

$$B_{r i} = \sum_{k=1}^n |r_{0i} - r_{0k}| \tag{18}$$

According to the calculation of Eqs. 16, 17, 18, the Comprehensive Grey Relation Degree of each minimum cut set of automobile doors *Expansion Adhesive Debonding* fault is shown in Table 8.

Table 8 Comprehensive Grey Relation Degree of minimum cut sets for *Expansion Adhesive Debonding* faults in automobile doors

Min cut set C_i	C_1	C_2	C_3	C_4	C_5	C_6
Total dispersion of Absolute Grey Relation Degree $B_{\varepsilon i}$	0.3031	0.2524	0.3316	0.2596	0.4039	0.3031
Total dispersion of Relative Grey Relation Degree $B_{r i}$	0.4910	0.2885	0.4910	0.6566	0.2480	0.2480
partition coefficient θ_i	0.3817	0.4666	0.4031	0.2833	0.6196	0.5500
Comprehensive Grey Relation Degree ρ_{0i}	0.6738	0.6576	0.6930	0.7089	0.6242	0.6364
Min cut set C_i	C_7	C_8	C_9	C_{10}	C_{11}	
Total dispersion of Absolute Grey Relation Degree $B_{\varepsilon i}$	0.5983	0.4342	0.4342	0.2596	0.2596	
Total dispersion of Relative Grey Relation Degree $B_{r i}$	0.2480	0.2480	0.2480	0.2480	0.7349	
partition coefficient θ_i	0.7070	0.6365	0.6365	0.5114	0.2610	
Comprehensive Grey Relation Degree ρ_{0i}	0.6036	0.6691	0.6691	0.6530	0.6137	

According to Table 8, we get the Importance Ranking of the minimum cut sets for *Expansion Adhesive Debonding* fault of automobile doors based on F-CGRA method, as:

$$C_4 > C_3 > C_1 > C_8 = C_9 > C_2 > C_{10} > C_6 > C_5 > C_{11} > C_7$$

Fig. 5 is minimum cut sets Comprehensive Grey Relation Degree histogram.

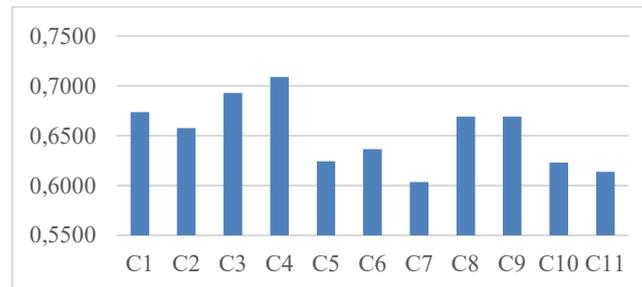


Fig. 5 Minimum cut sets Comprehensive Grey Relation Degree histogram

5. Result analysis and discussion

The FTA results of the three methods are compared. The results show that if a basic event occurs in a different minimum cut set, the probability of occurrence of the top event is greatly reduced after that event is excluded, even if the probability of occurrence of that event is low. This finding is also proved by the practice in the actual production. Unlike the FTA method alone, the combination of FTA and GRA can effectively analyse the relationship between system fault characteristics and intuitively reflect the likelihood magnitude of each basic event fault in the system.

In the actual production, the fault analysis team of G Group found the *Expansion Adhesive Debonding* fault in the automobile doors after 13 days through several checks and experiments, such as measurement of sheet metal parts, adjustment of the spacing between sheet metal parts, adjustment of glue application position and glue application amount, and high-temperature exposure of expansion adhesive placement, using inspection tools. The group found that the main cause of the problem is the intermediate events M_1 (the nature of expansion adhesive was not enough to meet the shipping demand) and M_3 (the spacing between the sheet metal parts was excessive), which include the basic events N_1-N_8 , $N_{11}-N_{13}$. After the improvement of the basic events N_5 , N_6 , N_7 , N_8 , N_{12} , N_{13} , N_{14} , the *Expansion Adhesive Debonding* fault of the automobile doors was prevented from occurring.

The FTA in the third part shows that the basic events N_5 , N_6 , N_7 , N_8 , N_{12} , N_{13} and N_{14} constitute the basic event combination $\{N_5, N_6\}$, $\{N_7, N_8\}$, $\{N_{12}\}$, $\{N_{13}\}$ and $\{N_{14}\}$ corresponding to the five failure modes C_8 , C_9 , C_3 , C_4 and C_{10} . This order is consistent with the F-AGRA method's ranking of fault levels in this study. $C_8 = C_9 > C_3 > C_4 = C_{10} > C_2 > C_{11} > C_1 = C_6 > C_5 > C_7$ indicates a high degree of conformity. Therefore, the F-AGRA method is the most consistent with the actual production situation. It is followed by the F-CGRA method and the F-RGRA method in the analysis of the *Expansion Adhesive Debonding* fault of automobile doors.

The comparison result of the three methods indicates that F-AGRA focuses on factors with a great influence on the system. The contribution of the failure modes represented by each minimum cut set to the top event is calculated. F-RGRA obtains the initial value of the vector set $X'_0 - X'_{11}$ by dividing the largest term in the vector $\max \{x_i\}$, which amplifies the influence of the failure mode changes in each minimum cut set on the top event. The degree of influence on the change of the top event is focused on reflecting the trend of factors and system changes. However, the direct influence of the failure mode represented by its minimum cut sets on the top event is not as obvious as F-AGRA. F-CGRA calculates the total deviation of the two correlations' balance reflects the direct contribution and variable influence of the fault mode represented by the minimum cut set to the top event.

For the direct contribution and change in the impact of the ranking, the unilateral accuracy and embodiment of the trend in the combined method are not as obvious as those in the two separate methods. Therefore, the method failure analysis between the above two methods is conducted.

The characteristics and applicable scenarios of the three methods are summarized in Table 9.

Table 9 Characteristics and applicable occasions of F-AGRA, F-RGRA, and F-CGRA methods

Method	Characteristic	Applicable occasions
F-AGRA	Visually reflect the contribution of the fault modes represented by each minimum cut set to the top event	The fault tree analysis is mainly composed of or gate structure events. and there are few and gate structure events, the interaction between elementary event is less
F-RGRA	The focus is on reflecting which factors are consistent with the system change trend, but the direct impact of the fault mode represented by the minimum cut set on the top event is not obvious	The fault tree analysis is mainly composed of and gate structure events, or there are few or gate structure events
F-CGRA	Balance reflects the direct contribution and changing impact of the fault mode represented by the minimum cut set on the top event	The fault tree analysis has many levels and is complex, and there are and gate events close to the root of the fault tree analysis

6. Conclusion

This study analyses the occasional and difficult-to-diagnose fault problems in the automotive industry. It uses three methods for fault analysis, explores the applicable scenarios of the three methods, improves the efficiency of fault analysis, and provides the automotive manufacturing industry with new ideas for fault analysis and fault diagnosis. Thus, the causes of product failure can be efficiently targeted for improvement, and the efficiency of improvement can be improved. This study can also be extended to the failure analysis of other complex products or systems by providing an efficient, accurate, and scientific theoretical basis for handling accident priorities, controlling the occurrence of accidents, and improving the reliability and safety of systems.

In this study, the method is only used to study the *Expansion Adhesive Debonding* fault of automobile doors. The method will be verified using other cases in the future. The analysis of the top event relies on the accurate construction of the fault tree and the comprehensive investigation of the cause of the failure. The probability of failure of each basic event must be accurately determined; otherwise, accurate conclusions may not be drawn. The fault diagnosis analysis method based on fault tree and generalized grey theory remains complicated in the calculation process when used for complex fault tree structures. Future research should focus on the sensitivity analysis of fault events while designing the program software using the three methods matching the operation to solve the problem of complex operation.

Acknowledgment

The authors gratefully acknowledge the support of Liaoning Province Education Department Project (NO.JYTMS20-230430: Research on environmental friendly mechanism of machinery and equipment remanufacturing process and industrial chain under the background of "double carbon") and Zhejiang Geely Holding Group Co., LTD. The authors would like to thank the editor and reviewers for their constructive suggestions of the paper.

References

- [1] Ortmeier, F., Schellhorn, G. (2007). Formal fault tree analysis – Practical experiences, *Electronic Notes in Theoretical Computer Science*, Vol. 185, 139-151, doi: [10.1016/j.entcs.2007.05.034](https://doi.org/10.1016/j.entcs.2007.05.034).
- [2] Ren, L.H., Hao, L.Q., Lin, L., Wang, H. (2019). Qualitative analysis of thermal disasters of belt conveyor based on fault tree method, *Chinese Journal of Grain and Oil*, Vol. 34, No. S2, 25-28.
- [3] Arcidiacono, G. (2003). Development of a FTA versus parts count method model: Comparative FTA, *Quality and Reliability Engineering International*, Vol. 19, No. 5, 411-424, doi: [10.1002/qre.537](https://doi.org/10.1002/qre.537).
- [4] Wang, X., Deng, J., Xiong, J., Xiao, Q., Chu, B., Liu, J. (2020). Construction and application of ICU sedative and analgesic management strategy based on Fault Tree Theory and eCASH (early comfort using analgesia, minimal sedatives and maximal humane care) concept, *Journal of Nursing*, Vol 35, No. 4, 39-43, doi: [10.3870/j.issn.1001-4152.2020.04.039](https://doi.org/10.3870/j.issn.1001-4152.2020.04.039).
- [5] Yazdi, M., Zarei, E. (2018). Uncertainty handling in the safety risk analysis: An integrated approach based on fuzzy fault tree analysis, *Journal of Failure Analysis and Prevention*, Vol. 18, No. 2, 392-404, doi: [10.1007/s11668-018-0421-9](https://doi.org/10.1007/s11668-018-0421-9).

- [6] Liu, J.Y., Leng, J.Q., Shang, P., Luo, L.J. (2022). Analysis of factors affecting the severity of highway accidents under ice and snow roads, *Journal of Harbin Institute of Technology*, Vol. 54, No. 3, 57-64.
- [7] Zhang, L.G., Liu, S.M., Li, K.F., Xu, J., Wang, S.N., Li, Q., Mei, Y.Y., Li, C.H., Yang, B. (2021). Prediction of shield construction risks in subway tunnelling based on fault tree and Bayesian network, *Modern Tunnel technology*, Vol. 58, No. 5, 21-29+55, doi: [10.13807/j.cnki.mtt.2021.05.003](https://doi.org/10.13807/j.cnki.mtt.2021.05.003).
- [8] Luo, C.K., Chen, Y.X., He, Z., Li, Y., Zhang, Y.M. (2021). Evaluation method for contribution rate of aviation equipment architecture based on fault tree analysis, *Journal of National University of Defense Technology*, Vol. 43, No. 1, 155-162, doi: [10.11887/j.cn.202101020](https://doi.org/10.11887/j.cn.202101020).
- [9] Liu, J.F., Li, Y.L., Ma, X.M., Wang, L., Li, J.L. (2021). Fault tree analysis using Bayesian optimization: A reliable and effective fault diagnosis approaches, *Journal of Failure Analysis and Prevention*, Vol. 21, 619-630, doi: [10.1007/S11668-020-01096-1](https://doi.org/10.1007/S11668-020-01096-1).
- [10] Shanks, K., Hamad, A., Ameer, A. (2020). Failure modes effects and criticality analysis and fault tree analysis case study of waste heat recovery plant in a cement factory, United Arab Emirates, *Journal of Failure Analysis and Prevention*, Vol. 20, 40-50, doi: [10.1007/s11668-020-00827-8](https://doi.org/10.1007/s11668-020-00827-8).
- [11] Lin, F.F., Huang, X., Huang, X.X. (2021). Quality evaluation of tinosporae radix from different producing areas by grey correlation analysis, *China Pharmacist*, Vol. 24, No. 10, 1942-1945, doi: [10.19962/j.cnki.issn1008-049X.2021.10.035](https://doi.org/10.19962/j.cnki.issn1008-049X.2021.10.035).
- [12] Dong, X.J., Chen, Y.W., Li, M., Zhang, Y.X. (2013). A spacecraft launch organizational reliability model based on CSF, *Quality and Reliability Engineering International*, Vol. 29, No. 7, 1042-1054, doi: [10.1002/qre.1455](https://doi.org/10.1002/qre.1455).
- [13] Liu, S.F., Xie, N.M. (2014). *Grey system theory and its application*, 4th Edition, China Science Press, Beijing, China, 66-83.
- [14] Pang, J.H., Dai, J.K., Zhou, H.Y., Li, Y. (2022). A new fault diagnosis method for quality control of electromagnet based on T-S fault tree and grey relation, *International Journal of Reliability, Quality and Safety Engineering*, Vol. 29, No. 4, Article No. 2141006, doi: [10.1142/S0218539321410060](https://doi.org/10.1142/S0218539321410060).
- [15] Tien, D.H., Van Bong, P., Hung, L.T. (2022). Applying improved fuzzy grey relation analysis algorithm in multi objective optimization for high-speed milling of 4Cr5MoSiV steel, *Process Integration and Optimization for Sustainability*, Vol. 6, No. 3, 587-601, doi: [10.1007/S41660-022-00238-2](https://doi.org/10.1007/S41660-022-00238-2).
- [16] Wang, S.B., Li, C.W., Sun, X.G. (2013). Search method for wind turbine gearbox failures based on grey and fuzzy fault tree analysis, *Journal of Applied Sciences*, Vol. 13, No. 6, 901-906, doi: [10.3923/jas.2013.901.906](https://doi.org/10.3923/jas.2013.901.906).
- [17] Wang, P.S. (2010). The application of grey relational fault tree analysis in mine hoist overrelaxation accident, *Applied Mechanics and Materials*, Vol. 33, 185-189, doi: [10.4028/www.scientific.net/AMM.33.185](https://doi.org/10.4028/www.scientific.net/AMM.33.185).
- [18] Zhang, G.J., Ye, Y.Q., Cui, T.J., Ma, Y.D., Wang, L.G. (2015). Fault mode analysis of tramcar falling by improving the generalized grey relational fault tree analysis, *Mathematical Practice and Understanding*, Vol. 45, No. 23, 169-178.
- [19] Chen, X., Gong, B.G., Jiang, P. (2019). Reliability diagnosis of LSSC for hazardous chemicals based on improved generalized grey correlation fault tree analysis, *China Science and Technology of Work Safety*, Vol. 15, No. 7, 149-155.
- [20] Wang, P.-S. (2012). Analysis of mine shaft falling car accident based on fault tree analysis, *Coal Mine Machinery*, Vol. 33, No. 7, 203-204, doi: [10.13436/j.mkjx.2012.07.066](https://doi.org/10.13436/j.mkjx.2012.07.066).
- [21] Yang, H.L., Huang, G.M. (2019). Fracture analysis of release piston based on the minimum cut set of fault tree analysis, *Naval Science and Technology*, Vol. 41, No. 1, 133-137.
- [22] Li, J.W., Qiao, J.G., Fu, X., Liu, X.-L. (2019). Failure analysis of anti-floating bolt system based on fuzzy fault tree analysis, *Journal of Safety and Environment*, Vol. 19, No. 4, 1128-1134, doi: [10.13637/j.issn.1009-6094.2019.04.004](https://doi.org/10.13637/j.issn.1009-6094.2019.04.004).
- [23] Yu, L., Liu, Y.Z., Guo, X.H., Luo, X., Wang, M., He, J. (2022). Distance analysis of railway tunnel emergency rescue stations based on fault tree analysis method, *Railway Standard Design*, Vol. 66, No. 2, 111-116, doi: [10.13238/j.issn.1004-2954.202012070002](https://doi.org/10.13238/j.issn.1004-2954.202012070002).
- [24] Yang, C.F., Zhuang, C., Sun, J.S., Yan, X.C. (2018). Multifactor analysis of road traffic accidents, *Chin. Journal of Chongqing Jiao Tong University (Natural Science Edition)*, Vol. 37, No. 4, 87-95.
- [25] Tian, X., Li, L., Zhao, L. (2020). Carbon emission reduction path and effectiveness evaluation of hybrid power grid based on GRA model, *Mathematical Practice and Understanding*, Vol. 50, No. 2, 130-140.

Optimizing smart manufacturing systems using digital twin

Ojstersek, R.^{a,*}, Javernik, A.^a, Buchmeister, B.^a

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

ABSTRACT

Presented paper investigates the application of digital twins for the optimisation of intelligent manufacturing systems and focuses on the comparison between simulation modelling results and real-world production conditions. A digital twin was created in the Simio software environment using a data-driven simulation model derived from a real-world production system. Running the digital twin in real time, which was displayed graphically, facilitated the analysis of key parameters, including the number of finished products, average flow time, workstation utilization and product quality. The discrepancies were attributed to the use of random distributions of input data in the dynamic digital twin, as opposed to the long-term measurements and averages in the real-world system. Despite the limitations in the case study, the results underline the financial justification and predictive capabilities of digital twins for optimising production systems. Real-time operation enables continuous evaluation and tracking of parameters and offers high benefits for intelligent production systems. The study emphasises the importance of accurate selection of input data and warns that even small deviations can lead to inaccurate results. Finally, the paper highlights the role of digital twins in optimising production systems and argues for careful consideration of input data. It highlights the importance of analysing real-world production systems and creating efficient simulation models as a basis for digital twin solutions. The results encourage extending the research to different types of production, from job shop to mass production, in order to obtain a comprehensive optimisation perspective.

ARTICLE INFO

Keywords:

Smart manufacturing;
Digital twin;
Optimisation;
Simulation modelling;
Simio;
Case study

*Corresponding author:

robert.ojstersek@um.si
(Ojstersek, R.)

Article history:

Received 19 July 2023
Revised 17 December 2023
Accepted 20 December 2023



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

1. Introduction

With the transition to Industry 4.0, digitization brings special challenges and concerns for production. One way to deal with these challenges is to use a digital twin. Digital Twin has emerged as a central paradigm in the field of smart manufacturing and Industry 4.0, with profound implications for production management, control frameworks and the broader landscape of modern manufacturing. The digital twin is not just a technological artefact, but also a multi-layered construct that connects physical entities with their virtual counterparts. It encompasses various dimensions, including its connotation, its reference model, its applications, and the associated research questions [1]. In the context of complex product assembly shop floors, a digital twin-based smart production management and control is emerging. This framework uses the concept of the digital twin to improve the agility and responsiveness of manufacturing processes. It provides a dynamic, real-time representation of the physical system that enables predictive maintenance, optimized resource allocation and adaptive production scheduling [2]. The interplay between digital twin and smart manufacturing are explored and their importance for transforming traditional manufacturing processes is highlighted [3]. This approach enables real-time monitoring, data-driven deci-

sion-making and improved operational efficiency [4]. The study focuses on optimising manufacturing takt time within a supply chain, with an emphasis on the strategic placement of a decoupling point. While the study provides a basis for improved operational efficiency, it is important to discuss the assumptions and generalisability to different production environments. Applications range from predictive maintenance to quality control, highlighting the versatility and transformative potential of digital twin-driven smart manufacturing [5]. The correlation and comparison between digital twins and cyber-physical systems (CPS) form a critical discourse. Both concepts have similarities in bridging the physical and digital worlds. While CPS emphasizes the integration of computational algorithms with physical processes, digital twins extend this integration to a comprehensive virtual representation [6]. This study looks at the central role of planning algorithms in smart manufacturing and addresses the associated challenges. The study not only highlights the advances in algorithms, but also emphasises the need for adaptive algorithms that can cope with dynamic production environments. Understanding their correlation helps to refine the conceptual framework for smart manufacturing [7]. The fusion of digital twin and big data [8] is explored in a 360-degree comparison. Big data analyses enrich the concept of the digital twin by incorporating extensive data sets into the virtual representation, thus enabling improved analyses and well-founded decisions. The synergy between the digital twin and big data serves as a cornerstone in the evolution towards smarter manufacturing processes [9]. In exploring the practical applications of digital twins in production logistics, this study presents a testbed that uses real-time location data [10]. Although the study is promising, it provides ideas for thought on the challenges of scalability and the need for further validation in complex production ecosystems [11]. By examining the enabling technologies, challenges and open research areas related to digital twins, a comprehensive understanding of the current landscape is achieved. From advanced sensor technologies to the intricacies of data interoperability, the path to realizing the full potential of digital twins is fraught with challenges. However, these challenges also present opportunities for further exploration and innovation that push the field into uncharted territories [12]. The focus on dynamics in manufacturing presents a new paradigm [13]. This reference highlights the transformative potential of the digital twin in redefining shop-floor operations by providing real-time insights and adaptive capabilities critical to managing the complexity of modern manufacturing [13]. By integrating the digital twin and improved bacterial foraging, job scheduling becomes more adaptable in dynamic production environments [14]. However, the limits of algorithmic scalability and generalisability of the proposed approach underline the need for further investigation [15]. Researchers look at the values digital twin brings to the modelling landscape, navigates the challenges faced in implementation, and identifies enablers that are critical to successful modelling [16]. Anticipating future developments broadens the discourse. This study [17] examines the enabling technologies and potential applications and provides a forward-looking perspective on industrial development beyond Industry 4.0. The comprehensive review provides an examination of the concepts, technologies, and industrial applications of the digital twin [18]. This synthesis enriches our understanding by providing a panoramic view of the historical development, current technological landscape, and practical applications of the digital twin in various industry sectors. By integrating the digital twin and improved bacterial foraging, job scheduling becomes more adaptable in dynamic production environments.

The digital twin plays a central role in the vision of the smart factory. It enables the transition from analysing the past to predicting the future. The concept of the digital twin represents a transformative force in the landscape of smart manufacturing and Industry 4.0. Its multidimensional nature, as explored through the systematic literature review and research, highlights its versatility and the need for further exploration. The intersection of digital twins with smart manufacturing frameworks, cyber-physical systems, big data, and their associated challenges points to a promising path for the future of manufacturing paradigms.

2. Methods

2.1 Digital twin

In the rapidly evolving landscape of modern manufacturing, the integration of digital technologies has revolutionised traditional approaches, giving rise to the concept of digital twins and paving the way for the era of smart manufacturing. At the centre of this transformative paradigm is the seamless connection between the physical system and its virtual counterpart, creating a symbiotic relationship that increases operational efficiency, productivity, and overall manufacturing system performance. The physical system, as the cornerstone of the production environment, embodies the tangible entities and processes that are essential to production. In parallel, sensing devices, consisting of sensors, internet of things devices and other data sources, play a central role by actively collecting real-time data from the physical system. This data, which is often voluminous and diverse, serves as the lifeblood of the digital twin ecosystem and requires a robust communication infrastructure for its transmission to the data processing unit. The communication infrastructure is an important link in the data flow chain that enables the transmission of the collected data to the data processing unit. This unit, which is responsible for pre-processing and organising the incoming data, plays a crucial role in ensuring data quality and relevance. The importance of this step cannot be overestimated, as the effectiveness of subsequent analyses and decision-making processes is highly dependent on the integrity and accuracy of the processed data. Subsequently, the focus is on data collection, in which the processed data is recorded and stored for further analysis. This stock of information forms the basis for the creation of the digital twin model. The digital twin, a virtual representation of the physical system, contains current states, behaviours and other relevant attributes derived from the collected data. In the field of intelligent manufacturing, the digital twin model goes beyond mere simulation and becomes an intelligent entity capable of making informed decisions [19]. Control systems capitalise on the insights offered by the digital twin by using it as a virtual mirror of the physical system, enabling precise decision-making and issuing commands to the control units [20]. This synergy between the digital and physical worlds creates a closed loop in which the digital twin continuously refines its understanding and responses based on feedback from the real world. As we delve into the scientific study of digital twins in smart manufacturing, the intricate interplay between the physical system, sensing devices, communication infrastructure, data processing, data acquisition, modelling of the digital twin and control systems underpinning this transformative paradigm becomes apparent. By deciphering the complexity and dynamics within this interconnected framework, we can unlock unimagined potential for efficiency, adaptability, and innovation in the manufacturing landscape, as shown in Fig. 1.

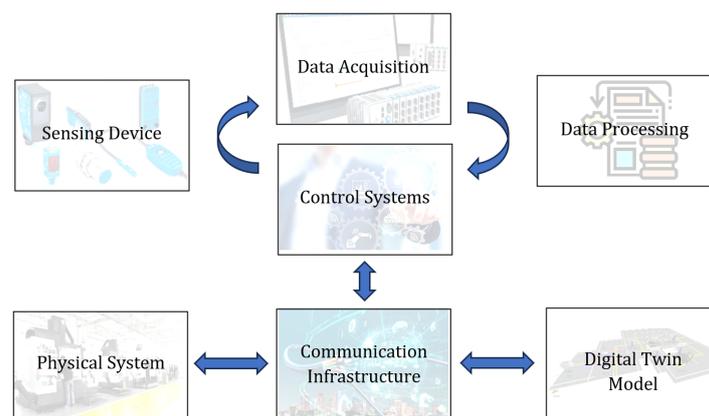


Fig. 1 Digital twin proposed block diagram structure

- **Physical system:** Represents the actual physical entity or process in the manufacturing environment.
- **Sensing Devices:** Collect data from the physical system. These can be sensors, IoT devices or other data sources.
- **Communication infrastructure:** Transmits the collected data to the data processing unit.

- **Data processing:** Includes the pre-processing and organisation of the data. This step is crucial to ensure the quality and relevance of the data.
- **Data acquisition:** Captures and stores the processed data for further analysis.
- **Digital twin model:** Uses the captured data to create a virtual representation of the physical system. This representation includes the current state, behaviour and other relevant attributes.
- **Control systems:** Use the digital twin model to make decisions and send commands to the control units.

2.2 Simio

Simulation modelling is a central aspect of decision-making processes that relies on simulation languages with different prerequisites for the programming language. This paper describes the Simio programming environment, which is characterised by its simulation language SIMIO, developed in the last decade. This language is characterised by the integration of 3D animations and graphical representations, which increase the visual appeal. The Simio software environment facilitates the creation and use of dynamic 3D animation models and offers a comprehensive library of standard elements. This library includes elements such as sources, sinks, servers, combiners, separators, resources, vehicles, workers, basic nodes, transfer nodes, connectors, paths, time paths and conveyors.

- **Source:** Serves as the entry point for new workpieces/orders and creates entities with specific types and arrival patterns.
- **Sink:** Represents the ending point for workpieces and is responsible for terminating entities that have completed processing.
- **Server:** Represents a workstation where processes are executed on workpieces.
- **Combiner:** Integrates individual entities, e.g. by assembling a product from different components.
- **Separator:** Works in the opposite way to a combiner by splitting one entity into several or creating several copies.
- **Resource:** Represents any intermittent resource that is required for a job and then returns to its intended location.
- **Vehicle:** Represents a transport mechanism that either has a predetermined route or is executed in response to specific requests.
- **Worker:** A mobile resource that is activated for specific tasks or to facilitate entity transfers between node locations.
- **Basic node:** Represents a connection point with no special functions, typically used upon arrival at the workstation.
- **Transfer node:** A dedicated node that is used to implement transport logic and is primarily used when leaving the workstation.
- **Connector:** Establishes connections between individual nodes; it is characterised by negligible time expenditure and is suitable for cases in which travel time is irrelevant.
- **Path:** Represents a classic transport path in which the user specifies both the distance and the speed of the element being traversed, thereby influencing the travel times.
- **Time path:** A special path in which the user specifies the duration of time that the element travels through.
- **Conveyor:** Represents a perpetual transport system, e.g. a constantly moving belt.

Each element is characterised by properties, states, events, appearance, and logic. Properties include user-defined input values, while states represent dynamic values that can change during the simulation. Events are temporal triggers that notify other elements of subsequent events. The appearance is displayed as a 3D graphical representation of the element within the model. Logic encapsulates a model within a model by prescribing the element's reactions to certain events and thus specifying its behavioural dynamics. Once the model is finalised, the Simio software environment offers the possibility to define different scenarios. These scenarios allow different param-

ters to be analysed simultaneously and their combined effects on the system to be observed. Individual scenarios facilitate the modification of model properties and provide insights into their isolated effects on the dynamics of the system in general.

2.3 Case study description

In the first phase of creating the simulation model, which was based on the architectural layout of the company's production facility in the Simio software environment, we meticulously described the spatial allocation of the machine installations while maintaining realistic dimensions. The existing transport and transit routes, warehouses and ancillary rooms were fully taken into account. It was decided to start the simulation model with an authentic workplace layout. The specific locations of the operational workstations were determined from the floor plan of the production hall and then integrated into our model using the dimensions initially determined. This preparatory step simplifies the positioning of the subsequent simulation elements, where several constraints were identified and carefully considered when creating the model:

- Element constraint: the specification that the model must not contain more than 30 elements required a pragmatic solution for the source and sink of entities, so these components had to be merged into a single source and sink.
- Time specification: The time frame of the simulation was limited to 10 working days. In order to adapt to this time parameter, the processing and setup / tear-down times were divided proportionally by 10 to ensure the temporal congruence of the simulation.

Before the actual construction of the model, overarching goals were defined to capture the planned attributes of the simulation:

- Different attributes: Different sequences of orders, processing times and setup and tear-down times depending on the type of entity created.
- Variable entity frequency: Entities of different types are generated with different frequencies.
- Type-dependent entity generation: The number of entities generated simultaneously depends on the type of entity generated.
- Trigger mechanism: The generation of new entities is triggered by the release of product A.
- Operational monitoring: The simulation interface provides a continuous overview of the length of the queues at the individual work centres.

In line with these constraints and the formulated objectives, three different entities were defined in the model, using the Model Entity element and symbolising the workpieces with the designation's product A, product B and product C. To increase the clarity of the model, each entity was clearly colour-coded. In addition, a travelling speed of 2 m/s was defined between the individual workstations, which represents product transport with forklift trucks. Assuming that an entity corresponds to a product batch, the intrinsic value of the entity was determined by multiplying the product cost and the batch size.

2.4 Digital twin model in Simio

The Simio simulation environment not only facilitates the creation of digital twins, but also offers a significant advantage when utilising input data from real-world production systems. The data-driven simulation model, complemented by a user-friendly graphical interface, not only enables efficient and robust numerical calculations, but also ensures a fast response time. This fast calculation capability is crucial for the successful implementation of a digital twin. In addition, the short calculation times in Simio enable the development of a digital twin where the numerical results can be seamlessly integrated into the real-time simulation within a visual virtual reality (VR) interface. This integration improves the overall experience of the simulation and provides a dynamic and interactive representation of the system behaviour. The use of a visual VR interface contributes to a deeper understanding of the simulated processes and facilitates in-depth analysis and decision making to optimise the real production system. The input data comes from a real-world production system consisting of seven workstations. Table 1 shows the main input data

that drives the data-driven simulation model and describes the financial, time, energy, location, and efficiency parameters of the real-world production system.

In the production system shown in Fig. 2, our analysis revolves around the dynamic flow of three different products that pass through a series of seven strategically positioned workstations. Each workstation is supervised by a dedicated employee who ensures that the production processes run smoothly. The intricate choreography of the production cycle involves the skilful use of a forklift to enable the smooth transport of raw materials at the beginning of the system, navigating through the intermediate stages and finally delivering the finished products. The simulation model ingeniously visualises the operation of the system by assuming a two-shift operation, with each shift lasting twelve hours, over the course of a conventional working week. This scheduling strategy is a fundamental part of the simulation as it provides a realistic representation of the daily operation of the system. Crucially, the efficiency of this production system is quantified through theoretical calculations of its production capacity. Based on these calculations, the system has a planned utilization rate of 85 %, which emphasises its ability to effectively convert input resources into products. This utilization metric serves as a valuable benchmark for evaluating and optimising the performance of the production system and provides insights for potential improvements and operational refinements.

Table 1 Digital twin input parameters

Workplace	Cut 1	Cut 2	Machining 1	Machining 2	Machining 3	Assembly 1	Assembly 2
Operating cost (EUR/h)	54	54	42	42	42	38	38
Idle cost (EUR/h)	36	36	21	21	21	13	13
x_{loc} (m)	9.6	9.6	15	22.2	29.4	36.6	43.2
y_{loc} (m)	11.4	5.4	1.2	1.2	1.2	1.2	6.6
Setup time (min)	12	12	18	18	18	9.6	9.6
Machine operation energy consumption (kWh)	12	12	8	8	8	1.5	1.5
Machine idle energy consumption (kWh)	1.5	1.5	1	1	1	0.2	0.2
Machine scrap (%)	2	1.5	4	4.5	5	1.5	2

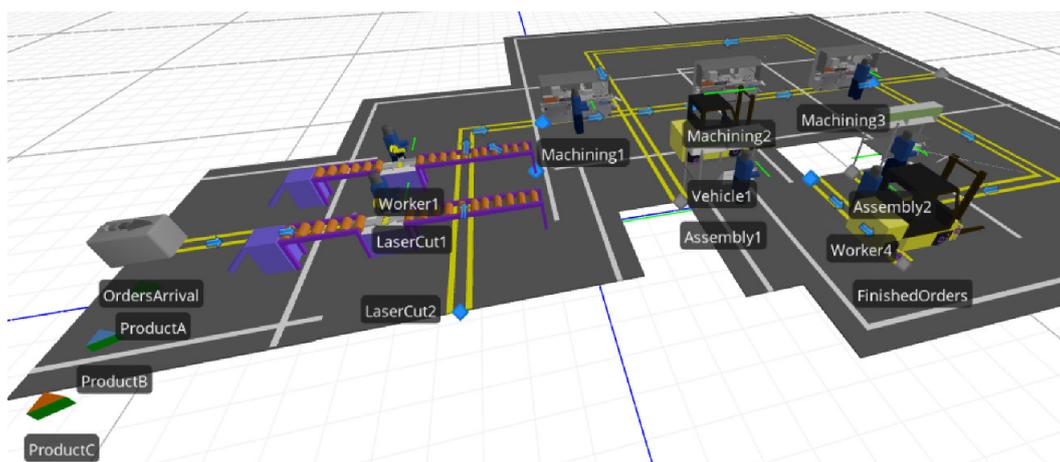


Fig. 2 Digital twin model in Simio

3. Results

The section dedicated to results is primarily concerned with highlighting the discrepancy between the results of a digital twin created in the Simio software environment and the results that emerge from the empirical reality of the real-world production system. The scientific literature emphasises the acceptance of results generated using traditional simulation methods. A discrepancy of

less than 3 % between the simulation results and the empirical reality is considered satisfactory. Conversely, any escalation beyond this threshold represents an increased risk of inefficiency of the simulation model, especially in terms of economic, safety and other operational constraints.

Table 2 presents the numerical results of the comparative analysis between the authentic production system and its digital counterpart, specifying the parameters for quantifying the final products and determining the average throughput intervals. The collected comparative results show deviations between the results of the digital twin and the real-world production environment. It is worth noting that a decadal average, derived from ten iterations of the simulation model and representing a normal five-day working week, shows a recognisable discrepancy in four out of seven parameters examined. For three parameters, however, the discrepancy remains below the defined threshold value of 3 %. The limited validity of the results emphasises the fact that the analysis embedded in the presented digital twin shortens its focus in time, while the input data of the authentic production system has a longer temporal extension.

Fig. 3 shows the graphical representation of the empirical production system contrasted with the digital twin, focussing on the parameter describing the quantity of finished products. The empirical evidence shows that the digital twin exactly fulfils the allocation rule for the proportional distribution of the workload depending on the composition of products A, B and C.

The number of completed products under the updated production system exceeds the corresponding yield of the digital twin. This discrepancy is due to the contrast between the use of stochastic distribution functions for the input parameters and the imposition of static values assumed by the real-world production system. The strategic integration of stochastic distribution functions into the simulation model of the digital twin serves to mimic authentic diurnal fluctuations inherent in the operating cadence of the empirical production system.

Table 2 Real-world vs digital twin numerical results comparison

	Real-world results	Digital twin results	Difference [%]
Number of finished products	163	158	-3.07
Number of finished product A	95	92	-3.16
Number of finished product B	46	46	0
Number of finished product C	22	20	-9.09
Average flow time product A (h)	171.67	175.90	+2.46
Average flow time product B (h)	149.80	152.70	+1.94
Average flow time product C (h)	104.70	122.16	+16.68

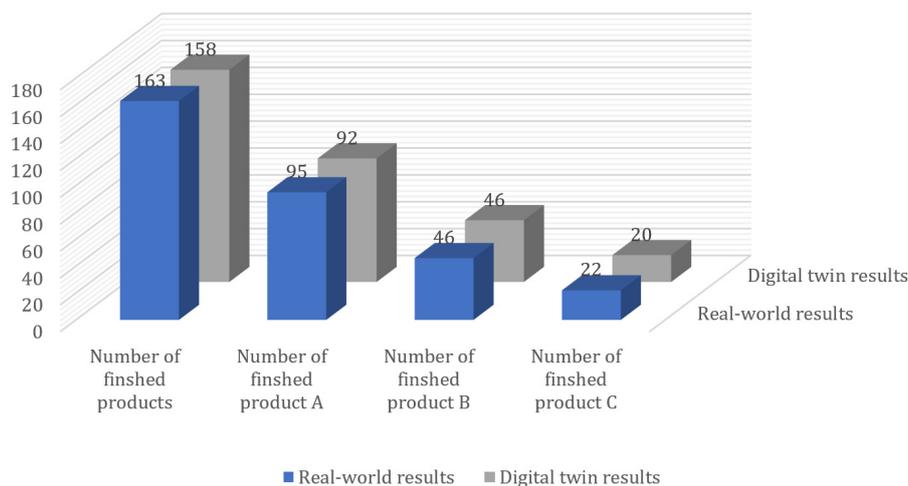


Fig. 3 Number of finished products comparisons

The parameter of an average flow time, as shown in Fig. 4, plays a central role in production system analysis. This graphical representation includes a quantitative representation of the average flow time taken during the entire production process. The average flow time calculated in this way is a key performance indicator that provides sophisticated insights into the complex dynamics that determine the efficiency, effectiveness, and overall vitality of the production system. This

metric is a fundamental benchmark that serves not only as an evaluation tool, but also as a catalyst for continuous improvement measures. Its importance stems from its ability to recognise the average flow time efficiency of the production system and to decipher crucial information about the progression of products or services over time from inception to completion. It also serves as a measurement tool for the systematic evaluation of the production process and provides a quantitative basis for identifying areas for optimisation. Average flow time plays a central role in the pursuit of the dual objectives of customer satisfaction and overarching corporate goals. By quantifying the time subtleties of production, it facilitates informed decision-making aimed at harmonising the production process with customer expectations and strategic business objectives. Consequently, the metric serves as a fundament for adaptive strategies that enable companies to proactively address challenges, optimise resource allocation and improve overall operational efficiency.

For the average flow time parameter, we find that the deviation between the results of the real-world production system and the result of the digital twin for products A and B is acceptable. However, for product C, where the production quantity is low, larger deviations occur. This is since a small number of products has a proportionally large influence on the deviation of the result comparison.

Table 3 and Fig. 5 show the comparative results of the workstations utilization and the percentage of well-made products. In this case, the comparative results illustrate the difference between the values of the real-world production system and the results of the digital twin. As the case under consideration is an order-related production, the results show that the system has a high-capacity utilization. The order planning and the utilisation of the available capacities are effective, so that there are only minor delays due to micro-interruptions and irregularities in the transport of raw materials and intermediate stocks. It is noteworthy that the utilisation in the real-world production system is lower than indicated by the digital twin. This emphasises the need for effective management in order-related production, which poses considerable challenges for the long-term sustainable operation of the production system due to its highly dynamic nature.

There are also differences in the parameter of the appropriate number of finished products; the results show that there are relatively higher values of insufficiently manufactured products in the real production system. Errors in production occur in individual operations, during transport and in the completion of the product. The discrepancy between the comparative results of the real-world production system and the digital twin can be attributed to the complexity of tracking and detailed monitoring of irregularities in the production/processing of semi-finished products within the digital twin, which occur continuously at all stages of production in the real-world production system.

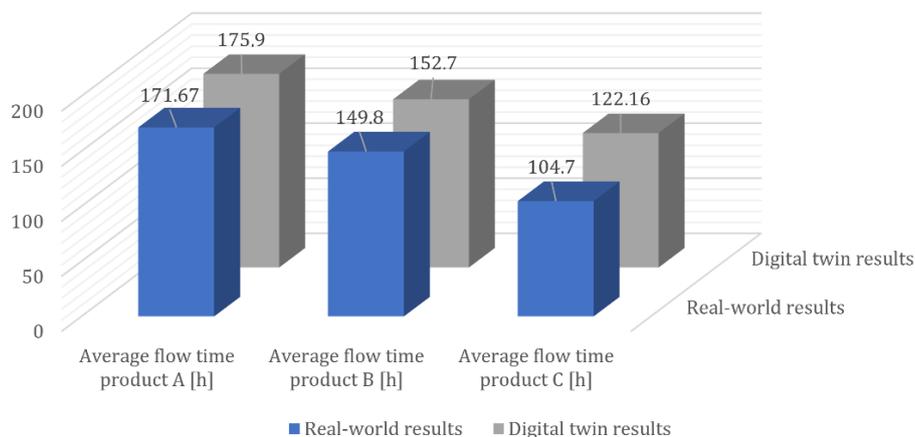


Fig. 4 Average flow time comparison

Table 3 Average utilization and quality products numerical results comparison

	Real-world results	Digital twin results	Difference (%)
Average utilization (%)	83.6	87.8	+4.2
Quality products (%)	91.5	95.3	+3.8

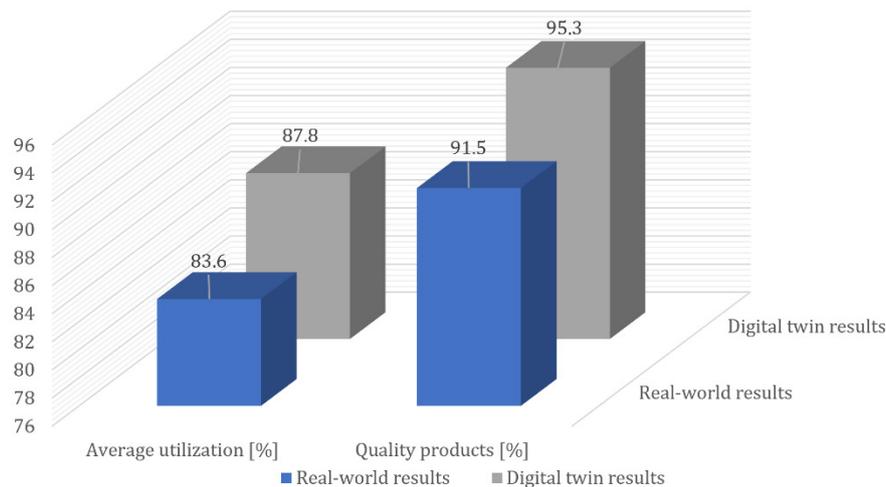


Fig. 5 Average utilization and quality products graphical presentation

4. Discussion

In line with the original research question of whether digital twins can be used to optimise production systems and whether simulation results are comparable with the real state of production, this article presents the construction of a digital twin within the Simio software environment. The digital twin is based on a data-driven simulation model of a real-world production system. The input data describes the digital twin in a complex way and enables its execution in real time through a graphical representation of the task execution in production. The numerical and graphical results obtained show the description of the production system in terms of parameters such as the number of finished products, the average flow time, the average utilization of the work centres and the appropriate quality of finished products.

In a comparative study of the results of the real-world production system and the digital twin, an average deviation of 4.9 % is found. This deviation exceeds the threshold value of 3 % recommended in the literature. From the point of view of exceeding the threshold, there are obvious overruns in parameters describing infrequently repeated measurements where individual measurements account for a significant common proportion. To address this, it would be advisable to extend the operation and calculation of the simulation model beyond the current five-day working week to a longer time frame, e.g., to a monthly or quarterly operation of the production system.

The observed discrepancies are due to the use of random distributions of input data by the digital twin, which is essential in highly dynamic production environments, in contrast to the real-world production system, where the parameters are measured over a longer period and the data represent averages of months. The research presented is limited by the case study considered, which indicates the need for extension and, in certain aspects, generalisation of the study. The results shown within the given limitations show that the implementation of digital twins enables the analysis of the performance of the production system with the aim of optimisation [21]. The main advantage of a digital twin in this regard is its financial justification, as it allows us to predict in advance, what will happen in the production system. The operation of a digital twin is also possible in real time during the continuous operation of the production system to evaluate and track certain parameters of the production system.

The presented approach shows the high benefit of digital twins in intelligent production systems. Given the high market dynamics in production systems, continuous optimisation is encouraged, and digital twins enable this by allowing the investigation of all important parameters in terms of efficiency, finances, environment, etc. When using digital twins, the results show that the right choice of input data is crucial, which must accurately represent the production system being analysed. Even a slight deviation in the data can lead to completely wrong results or a major deviation from the real-world situation. In this case, detailed collection, analysis, filtering and understanding of the input data is of utmost importance. Only in this way we can ensure the accuracy of the numerical and graphical results, whereby the final evaluation of the constructed digital twin also plays a decisive role.

5. Conclusion

In this paper, we explore the use of digital twins as an optimisation method for intelligent manufacturing systems. By using the digital twin, together with intelligent algorithms, companies can achieve data monitoring, improve the operation of the production system, and develop innovative products and services. We emphasise the importance of both studying the real-world production system and creating an efficient simulation model as a basis for a digital twin solution. The results show that digital twins can be used as a basis for the optimisation of production systems where the input data is one of the most important factors.

The experiments conducted were designed in a laboratory environment using input parameters from a real-world production system. In order to fully utilise the potential of the proposed digital twin, the scope of the studied problem needs to be extended to different types of production systems, from job shop to mass production. The proposed concept offers production companies a more comprehensive optimisation perspective for their production systems with the aim of increasing overall efficiency and global competitiveness. Despite the laboratory environment, this work emphasises the importance of using digital twin solutions in real-world production systems. In the current research landscape, our work represents a cutting-edge contribution to the investigation of the possibilities of the digital twin in intelligent production systems. The growing presence of the digital twin in the real world emphasises the importance of our findings.

This study provides new and fascinating insights into the field of digital twin methods as an optimisation tool for intelligent manufacturing systems. Our study lays the foundation for future research and recognises that the results presented here only capture certain findings that cannot be conclusively refuted. To increase the authenticity of our findings, we plan to extend our experiment in the future by expanding the types of manufacturing systems and investigating the usefulness of digital twins depending on the characteristics of the manufacturing systems. In further research steps, we plan to transfer the results of the digital twins to a real-world application where a comparison of real-time optimisation can be performed.

The implications of the presented research go beyond the presented framework and offer potential avenues for further investigation and practical applications in the field of digital twins and smart manufacturing. Through a comprehensive analysis of several parameters (average flow time, total number of finished products, average utilisation of workstations and quality of products), this paper contributes to the current understanding of the optimisation possibilities of smart manufacturing systems and lays the foundation for future research.

Acknowledgement

The authors gratefully acknowledge the support of the Slovenian Research Agency (ARIS), Research Core Funding No. P2-0190. The authors acknowledge the use of research equipment system for development and testing cognitive production approaches in Industry 4.0: Collaborative robots with equipment and Sensors, hardware and software for ergonomic analysis of a collaborative workplace, procured within the project "Upgrading national research infrastructures – RIUM", which was cofinanced by the Republic of Slovenia, the Ministry of Education, Science and Sport and the European Union from the European Regional Development Fund.

References

- [1] Jones, D., Snider, C., Nassehi, A., Yon, J., Hicks, B. (2020). Characterising the digital twin: A systematic literature review, *CIRP Journal of Manufacturing Science and Technology*, Vol. 29, Part A, 36-52, doi: [10.1016/j.cirpj.2020.02.002](https://doi.org/10.1016/j.cirpj.2020.02.002).
- [2] Zhuang, C., Liu, J., Xiong, H. (2018). Digital twin-based smart production management and control framework for the complex product assembly shop-floor, *The International Journal of Advanced Manufacturing Technology*, Vol. 96, 1149-1163, doi: [10.1007/s00170-018-1617-6](https://doi.org/10.1007/s00170-018-1617-6).
- [3] Lu, Y., Liu, C., Wang, K.I.-K., Huang, H., Xu, X. (2020). Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues, *Robotics and Computer-Integrated Manufacturing*, Vol. 61, Article No. 101837, doi: [10.1016/j.rcim.2019.101837](https://doi.org/10.1016/j.rcim.2019.101837).
- [4] Pattanaik, L.N. (2021). Simulation optimization of manufacturing takt time for a leagile supply chain with a decoupling point, *International Journal of Industrial Engineering and Management*, Vol. 12, No. 2, 102-114, doi: [10.24867/IJIEEM-2021-2-280](https://doi.org/10.24867/IJIEEM-2021-2-280).

- [5] Tao, F., Qi, Q., Liu, A., Kusiak, A. (2018). Data-driven smart manufacturing, *Journal of Manufacturing Systems*, Vol. 48, Part C, 157-169, [doi: 10.1016/j.jmsy.2018.01.006](https://doi.org/10.1016/j.jmsy.2018.01.006).
- [6] Workneh, A.D., Gmira, M. (2022). Scheduling algorithms: Challenges towards smart manufacturing, *International Journal of Electrical and Computer Engineering Systems*, Vol. 13, No. 7, 587-600, [doi: 10.32985/ijeces.13.7.11](https://doi.org/10.32985/ijeces.13.7.11).
- [7] Tao, F., Qi, Q., Wang, L., Nee, A.Y.C. (2019). Digital twins and cyber-physical systems toward smart manufacturing and industry 4.0: Correlation and comparison, *Engineering*, Vol. 5, No. 4, 653-661, [doi: 10.1016/j.eng.2019.01.014](https://doi.org/10.1016/j.eng.2019.01.014).
- [8] Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., Sui, F. (2018). Digital twin-driven product design, manufacturing and service with big data, *The International Journal of Advanced Manufacturing Technology*, Vol. 94, 3563-3576, [doi: 10.1007/S00170-017-0233-1](https://doi.org/10.1007/S00170-017-0233-1).
- [9] Qi, Q., Tao, F. (2018). Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison, *IEEE Access*, Vol. 6, 3585-3593, [doi: 10.1109/ACCESS.2018.2793265](https://doi.org/10.1109/ACCESS.2018.2793265).
- [10] Ojstersek, R., Buchmeister, B., Herzog Vujica, N. (2020). Use of data-driven simulation modeling and visual computing methods for workplace evaluation, *Applied Sciences*, Vol. 10, No. 20, Article No. 7037, [doi: 10.3390/app10207037](https://doi.org/10.3390/app10207037).
- [11] Hauge, J.B., Zafarzadeh, M., Jeong, Y., Li, Y., Khilji, W.A., Larsen, C., Wiktorsson, M. (2021). Digital twin testbed and practical applications in production logistics with real-time location data, *International Journal of Industrial Engineering and Management*, Vol. 12, No. 2, 129-140, [doi: 10.24867/IJIEEM-2021-2-282](https://doi.org/10.24867/IJIEEM-2021-2-282).
- [12] Fuller, A., Fan, Z., Day, C., Barlow, C. (2020). Digital twin: Enabling technologies, challenges and open research, *IEEE Access*, Vol. 8, 108952-108971, [doi: 10.1109/ACCESS.2020.2998358](https://doi.org/10.1109/ACCESS.2020.2998358).
- [13] Tao, F., Zhang, M. (2017). Digital twin shop-floor: A new shop-floor paradigm towards smart manufacturing, *IEEE Access*, Vol. 5, 20418-20427, [doi: 10.1109/ACCESS.2017.2756069](https://doi.org/10.1109/ACCESS.2017.2756069).
- [14] Ojstersek, R., Javernik, A., Buchmeister, B. (2022). Importance of sustainable collaborative workplaces - simulation modelling approach, *International Journal of Simulation Modelling*, Vol. 21, No. 4, 627-638, [doi: 10.2507/IJSIMM21-4-623](https://doi.org/10.2507/IJSIMM21-4-623).
- [15] Rasheed, A., San, O., Kvamsdal, T. (2020). Digital twin: Values, challenges and enablers from a modeling perspective, *IEEE Access*, Vol. 8, 21980-22012, [doi: 10.1109/ACCESS.2020.2970143](https://doi.org/10.1109/ACCESS.2020.2970143).
- [16] Maddikunta, P.K.R., Pham, Q.-V., Prabadevi, B., Depa, N., Dev, K., Gadekallu, T.R., Ruby, R., Liyanage, M. (2022). Industry 5.0: A survey on enabling technologies and potential applications, *Journal of Industrial Information Integration*, Vol. 26, Article No. 100257, [doi: 10.1016/j.jii.2021.100257](https://doi.org/10.1016/j.jii.2021.100257).
- [17] Liu, M., Fang, S., Dong, H., Xu, C. (2021). Review of digital twin about concepts, technologies, and industrial applications, *Journal of Manufacturing Systems*, Vol. 58, Part B, 346-361, [doi: 10.1016/j.jmsy.2020.06.017](https://doi.org/10.1016/j.jmsy.2020.06.017).
- [18] Huo, L., Wang, J.Y. (2022). Flexible job shop scheduling based on digital twin and improved bacterial foraging, *International Journal of Simulation Modelling*, Vol. 21, No. 3, 525-536, [doi: 10.2507/IJSIMM21-3-C014](https://doi.org/10.2507/IJSIMM21-3-C014).
- [19] Trebuna, P., Pekarcikova, M., Edl, M. (2019). Digital value stream mapping using the tecnomatix plant simulation software, *International Journal of Simulation Modelling*, Vol. 18, No. 1, 19-32, [doi: 10.2507/IJSIMM18\(1\)455](https://doi.org/10.2507/IJSIMM18(1)455).
- [20] Marc, I., Kušar, J., Berlec, T. (2022). Decision-making techniques of the consumer behaviour optimisation of the product own price, *Applied Sciences*, Vol. 12, No. 4, Article No. 2176, [doi: 10.3390/app12042176](https://doi.org/10.3390/app12042176).
- [21] Passath, T., Huber, C., Kohl, L., Biedermann, H., Ansari, F. (2021). A knowledge-based digital lifecycle-oriented asset optimisation, *Tehnički Vjesnik – Technical Gazette*, Vol. 15, No. 2, 226-234, [doi: 10.31803/tg-20210504111400](https://doi.org/10.31803/tg-20210504111400).

Incentive modeling analysis in engineering applications and projects with stochastic duration time

Zhao, J.^{a,*}, Su, J.F.^{b,*}

^aSchool of Management and Economics, University of Electronic Science and Technology of China, Chengdu, P.R. China

^bInternational College, Krirk University, Bangkok, Thailand

ABSTRACT

Incentives are quite common to be utilized in engineering applications such as some infrastructure development projects or construction projects. Due to the increasing complexity of construction management and the continuing trend towards outsourcing of component or engineering outsourcing activities, we focus on the issue of incentive design. Time collaboration is one of the main focuses of random project duration time in parallel projects. In this article, we consider a setting where a manufacturer outsources two parallel subtasks to two different suppliers, and the manufacturer is time sensitive. On the premise that the project completion time follows the exponential distribution, some models are established to compare the proposed incentives and we get the comparative analysis of the proposed incentives. This paper puts forward three kinds of time-based incentive mechanisms, namely, deadline incentive mechanism, competition mechanism and mixed incentive mechanism. We do modeling analysis for all incentive mechanisms. We get the optimal work rates determined by suppliers and compare various incentive mechanisms to maximize manufacturers' profits.

ARTICLE INFO

Keywords:

Engineering applications;
Incentive mechanism;
Modeling analysis;
Stochastic duration time;
Deadline incentive model;
Competition model;
Parallel subtasks;
Mixed model

*Corresponding author:

15982405709@163.com
(Zhao, J.)
jiafu.su@hotmail.com
(Su, J.F.)

Article history:

Received 12 July 2023
Revised 13 December 2023
Accepted 15 December 2023



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

1. Introduction

In recent years, with the increasing complexity of cross-industry projects, scholars have paid more and more attention to the importance and significance of engineering and project management research. In traditional construction or modern product development management, there is an increasing tendency to explore the role of incentive mechanisms in different environments. At the same time, project management and engineering applications have expanded dramatically from traditional building and infrastructure management to include new product development, information technology, pharmaceuticals, and service development. Project-based organizations are increasingly dependent on suppliers, and as a result, supply chain risks can arise. These new applications create different characteristics, and we need to develop new and innovative ways to solve the problems that may arise in these areas.

Project management includes planning and controlling the uncertainty in the duration of various components or tasks. Tools such as CPM, PERT and cost-time trade analysis are often used for project management. Therefore, the project owner can have a detailed overview when managing projects. These tools are effective and efficient when there is little uncertainty about the project completion time or operating costs. However, uncertainty arises in industries such as architecture, hardware and software development, national defense and management due to the increasing complexity of project management. For example, new power station, aircraft, and telecommunications network are all characterized by long construction period, high risks, and key links with engineering and commercial teams. We know little about how to deal with and manage the risks arising from project uncertainty.

In the field of consumer electronics, for example, more and more enterprises cooperate with Foxconn, Flextronics and other electronic manufacturing service providers to conduct research, development, design and production of electronic products. In a setting where the manufacturer cannot begin to work until all parts are delivered, the manufacturer's schedule depends on the slowest delivery. Project delivery delays can have a significant negative impact on the interests of manufacturers, so manufacturers need to devise mechanisms to encourage suppliers to work faster so as to speed up project completion time.

This article considers a setting where a manufacturer outsources two parallel subtasks to two suppliers with random project completion times. The manufacturer (she) plays the role of Stackelberg leader, making decisions about the best price she can offer to the suppliers. Each supplier (he) must determine the best work rate to obtain the best expected discount profit. When the manufacturer is time urgent to launch new products, she is very sensitive to the completion time of the entire project. We assume that the supplier's work rate is not adjustable during the whole period due to some objective factors or external conditions. We mainly focus on time-based incentive mechanisms in this paper.

Time-based incentive mechanism is common in practice and is a tool for manufacturers to motivate suppliers to complete tasks on time while managing random projects. Construction projects involve complex processes, so avoiding delays during the construction phase is very important. For example, after the Northridge earthquake in 1994, the city of Los Angeles (the project manager) wanted to repair the Santa Monica Freeway within 180 days. The city of Los Angeles quickly offered Clint Meyers (the contractor) an incentive contract. Specifically, if the project is completed 180 days ahead of schedule, the contractor will be paid an additional \$200,000 per day; However, a penalty of \$200,000 per day will be paid if the contractor is late [1, 2]. We propose three types of time-based incentives in our article.

In the case of deadline incentive mechanism, the project contractor is usually given a deadline in traditional project management and is responsible for the penalty charged by the project manager if the deadline is not met. The project owner often provides incentives/disincentives to the contractors. Deadlines are also commonly used in product development processes [3, 4]. In an empirical study of the global computer industry, Eisenhardt and Tabrizi [5] found that frequent deadlines help accelerate product development.

Due to the randomness of the task completion time, the two suppliers may end the tasks with different task duration times. The result could be that one supplier finishes earlier (hereinafter referred as faster supplier) and one later (hereinafter referred as slower supplier), and this asynchronization may harm the manufacturer's profit. Thus, we propose a competition incentive mechanism, which means that the faster supplier will be awarded, and the slower supplier will be punished.

On the problem of designing incentives for parallel random subprojects, our paper is the first one to integrate reward and punishment incentives design together with project structure. We also incorporate the criteria for rewards and penalties in those different incentive mechanisms, and we do a more comprehensive comparative analysis of proposed incentive mechanisms. We then combine the above two incentive mechanisms together, thus getting a mixed incentive mechanism. We then examine its influence on manufacturer's benefits. By comparing those three different incentive mechanisms, we get some illuminating results to shed light on project management and our analyses solve the following questions:

Q1: What is the supplier's optimal work rate under different incentive mechanisms?

Q2: What incentives will result in higher profits for the manufacturer?

The remainder of this paper is organized as follows: in Section 2, we provide a related literature review on project supply chain management. Section 3 presents our base model settings. In Section 4, we compare three kinds of incentive mechanisms together with the benchmark case. In Section 5, by some numerical analysis, we give manufacturer's optimal choice of incentive mechanisms under different conditions. In Section 6, we summarize our work and give a conclusion of this paper. All proofs are relegated to Appendix.

2. Literature review

Scholars and practitioners have extensively studied project management. In traditional project management, scheduling, and planning activities are the main concerns, and optimization is the main way to solve these problems. We recommend that readers refer to the study [6, 7] for further information. Hall [8] provides a comprehensive review and discussion of research and teaching opportunities in this area. In this section, we mainly focus on the problem of project contracting and incentive mechanism design within multi-agents. We review some important contributions that researchers have achieved. This paper mainly involves the following four research directions.

2.1 Deadlines in project

Deadlines are often used in product development processes and construction projects. A deadline means that the project must be stopped when the timeline reaches the stop point, or agents who do not meet the deadline will be penalized. Zhang [9] examines the value of deadlines from the agency theory perspective and considers that a company pays an agent to lead product development activities. It mainly focuses on the trade-off between the project's return and the project's labor cost. The paper concludes that deadlines are good, because they diminish the agent's incentive to procrastinate dynamically. Bordley *et al.* [10] consider a deadline problem with uncertainty. Traditional project management deals with unexpected changes in project deadlines through an external change control process. They introduce a new counter-intuitive idea of recognizing uncertainty in project deadlines and show that it can greatly enhance the value of managers' decisions. Du *et al.* [11] studied the optimal timing for occasional placement of 'fulcrum' in crowdfunding. The project will be funded successfully only if the target is met within the specified deadline. They evaluated three strategies in detail, namely, seeding, feature upgrading and time-limited offering, to increase the likelihood of turning the random pledge process from failure to success.

Our paper differs from the study [9, 11], in which the deadline serves as a stopping point of the project. However, ours mainly focuses on the reward or penalty associated with the specific deadline.

2.2 Time-based incentive contract

In the traditional supply contract, the uncertainty of demand is the main focus, and the quantity of order is the main decision variable. However, in project contracting, time uncertainty is the main focus, and the contractor's work rate decision is the main decision variable. Therefore, various incentive contracts based on time are studied by scholars in the field of project management. Gupta *et al.* [12] studied A+B infrastructure procurement mechanism, which state transportation agencies use to provide incentives for faster completion. Early completion will be rewarded, late will be penalized. Managers use time-based incentive contracts to encourage contractors to make greater efforts to complete tasks faster. Tang *et al.* [13] compared two different time-dependent project management contracts (C1 and C2) when managers conducted reverse auctions. In each incentive contract, completion earlier than the due date will be incentivized and completion later than the due date will be discouraged.

Chen and Lee [14] consider a coordination problem where the manufacturer carries out a series of tasks, and each task needs to procure a certain material from a supplier. They propose a delivery-schedule-based incentive contract to mitigate incentive misalignment between firms in a project supply chain. The delivery delay than the due date will be punished and early delivery will be awarded. Bayiz and Corbett [2] provide a framework to integrate contracting and asymmetric information into a project management context and study the value of incentive contracts. They derived the optimal incentive contracts the project manager offers to the subcontractors when the two subprojects are conducted in parallel or in serial, respectively. Kwon *et al.* [15] consider a coordination problem when managing a project with uncertain completion and an unobservable contractor's work rate. Fixed price, time-based, cost-based, and rate-based incentive contracts are examined in their work. Zhao and Mu [16] consider a situation where a manufacturer outsources two parallel tasks to two different suppliers, and they propose nine kinds of time-based incentive contracts a manufacturer can offer to the suppliers.

2.3 Reward-penalty mechanism

A project usually consists of a set of activities that introduce delays or expeditions related to penalty or reward issues when the planned completion time is incorrectly estimated. Bergantiños and Lorenzo [17] consider a situation where a planner wants to execute a project involving several companies. If the deadline is missed, the company will be penalized. They discussed two ways to impose penalties: they would apply only if the entire project was delayed; Even if the project is completed on time, the penalties apply to every company that causes delays. Estévez-Fernández [18] analyzed the situation where a project with multiple activities was not realized as planned. If the project is accelerated, there is a reward. Similarly, if the project is delayed, there is a penalty. In this paper, we consider the existence of any non-decreasing reward function and penalty function for total exploration and delay, respectively, focusing on the distribution of total reward (penalty) function between activities. Chen *et al.* [19] proposed an incentive payment contract for a series of random items. Their proposed contract reflects the convex time-cost trade-off that is well known in the project literature. In contrast to fixed-price contracts, such incentive contracts imply penalties for suppliers.

2.4 Parallel projects

Project structure is a key focus when conducting product development activities or outsourcing activities. These sub-projects are usually performed in parallel (tasks have to be performed in parallel), or in serial (tasks have to be performed sequentially) or in network (tasks have to be performed in parallel-serial). Our work mainly considers a scenario where the sub-projects are conducted in parallel. Kwon *et al.* [20] studied a delayed payment scheme where multiple concurrent subprojects were outsourced to different suppliers. Under the delayed payment scheme, each supplier will be paid after all tasks are completed, that is, faster supplier payments are delayed. Song *et al.* [21] explored incentives for firms under risk-sharing partnerships in the context of project management. In this partnership, each partner pays its own costs and shares project completion results (rewards or losses) between faster and slower suppliers. They examine the project network in serial, parallel, and assembled ways. Dawande *et al.* [22] studied the coordination problems faced by enterprises when undertaking projects consisting of multiple tasks. Under the assumption of exponential completion time, the optimal coordination contract of parallel and sequential tasks is studied.

The above literatures only consider the incentive mechanism design problem under partial factors. However, our paper considers the time-based incentive mechanism design problem faced by a manufacturer who carries out outsourcing activities in more complicated cases. Our proposed incentive mechanisms take deadlines in projects issue and reward-penalty incentives into consideration, and we also consider the time synergy problem in parallel projects. Our paper is a more comprehensive one that considers the incentive mechanisms design issue by the project owner.

3. Model settings

In our context, we consider a scenario where the manufacturer outsources two parallel sub-tasks with stochastic project completion time to two different suppliers. The duration of each subtask X_i ($i = 1, 2$) is exponentially distributed with parameter μ_i , where the work rate $\mu_i > 0$ is selected by supplier i at the beginning of the project. Throughout our analysis, we assume that the two subtasks are of the same workloads and difficulties. The suppliers cannot change the work rates once selected at the beginning, due to some technical or practical reasons. The manufacturer cannot start to work until the suppliers deliver all parts.

We consider the time value of revenue gained and cost incurred by the supplies into our context. Thus, our model takes the discounting issue into our analysis. Let the continuous-time discount rate be $\alpha > 0$, which captures the fact that the suppliers want to receive his payment earlier and incur costs later.

We consider a situation where the manufacturer is time sensitive to the completion time of the total project. For example, while developing some new products, the manufacturer is urgent to launch or release new products to occupy the market earlier. The total completion time of the project satisfies that $T = \max(X_1, X_2)$. For ease of exposition, we assume that the project's value $V(T)$ to the manufacturer is a linear, decreasing function of the project completion time, specifically, we define $V(T) = A - BT$ ($A, B > 0$).

By the properties of exponential distributions, we have that the probability density function of random variable X_i ($i = 1, 2$) is $\mu_i e^{-\mu_i t}$ and the cumulative distribution function of a random variable X_i is $F_i(t) = 1 - e^{-\mu_i t}$. We can infer that the discount factor satisfies that $E[e^{-\alpha X_i}] = \int_0^\infty \mu_i e^{-\alpha t} e^{-\mu_i t} dt = \frac{\mu_i}{\alpha + \mu_i}$.

While working at the rate of μ_i ($i = 1, 2$), each supplier incurs an operating cost of $\kappa(\mu_i)$ per unit time. We assume that the suppliers' operating cost $\kappa(\mu_i)$ per unit time associating with the work rate μ_i is a convex-increasing function, and it is given by $\kappa(\mu_i) = k\mu_i^2$ with $k > 0$. Furthermore, we have that supplier i 's expected discounted operating costs equal that $E\left[\int_0^{T_i} \kappa(\mu_i) e^{-\alpha t} dt\right] = \int_0^\infty \left[\int_0^{t_i} \kappa(\mu_i) e^{-\alpha t} dt\right] \mu_i e^{-\mu_i t_i} dt_i = \frac{k\mu_i^2}{\alpha + \mu_i}$.

We compare four kinds of mechanisms manufacturer can offer to the suppliers (three timed-based incentive models together with the benchmark case). For the sake of perceptual intuition, we list some characteristics among the four mechanisms. Table 1 shows the specific elements contained in these four kinds of models.

In the base model, we consider a setting where the manufacturer does not offer any incentive to the suppliers. The two suppliers start to work simultaneously and make the work rate decisions μ_i ($i = 1, 2$). Each supplier can get a payment of ω after his own subtask is finished, and the two suppliers make their decision independently.

In the deadline incentive model, the manufacturer imposes a specific due date, which we also referred as a deadline D . In addition, to paying a base price ω , the manufacturer awards the supplier a reward at r per unit time of early completion than the deadline, and charges the supplier a penalty at p per unit time of delay than the deadline.

In the competition incentive model, the manufacturer awards the faster supplier a reward of r per unit time earlier than the slower supplier; she also charges the slower supplier a penalty of p per unit time later than the faster supplier, together with a base price ω .

Table 1 Classification of different mechanisms

Base Model • A base price	Deadline Incentive Model • A base price • A specific due date. • An incentive/disincentive scheme
Competition Model • A base price • A comparison of the faster and slower • An incentive/disincentive scheme	Mixed Model • A base price • A specific due date. • A comparison of the faster and slower • An incentive/disincentive scheme

In the mixed incentive model, the manufacturer imposes a specific due date D . The manufacturer awards the faster supplier a reward of r per unit time earlier than the deadline and charges the slower supplier a penalty of p per unit time later than the deadline, together with a base price ω .

4. The model

4.1 Base model

In our base model, there is no incentive mechanisms offered to the suppliers by the manufacturer. The manufacturer outsources two subtasks to two different suppliers. The two suppliers start to work simultaneously, and each supplier does not take the other's decision into his own consideration. Thus, each supplier makes his work rate decision μ_i independently. When supplier i finishes his own subtask, he will get a payment of ω , and his expected discounted profit can be expressed as:

$$\Pi_i = \omega E[e^{-\alpha X_i}] - \frac{k\mu_i^2}{\alpha + \mu_i} \quad (i = 1, 2) \tag{1}$$

From the concavity of Eq. 1, we can get:

Proposition 1. Both suppliers' optimal work rates are the same in equilibrium, and they are characterized by

$$\mu^* = -\alpha + \sqrt{\alpha^2 + \frac{\omega\alpha}{k}} \tag{2}$$

We have that μ^* is strictly increasing with ω .

When both subtasks are finished and delivered to the manufacturer, the outsourcing activity ends. Because we have that $E[T] = E[\max(X_1, X_2)] = \frac{1}{\mu_i} + \frac{1}{\mu_{-i}} - \frac{1}{\mu_i + \mu_{-i}}$, thus, in equilibrium, we can get that $E[T] = \frac{3}{2\mu^*}$.

The manufacturer makes an optimal price decision ω , and her profit can be expressed as:

$$\Pi_m = A - BE[T] - 2\omega. \tag{3}$$

By the concavity of Eq. 3, we can obtain:

Proposition 2. The manufacturer's optimal price decision satisfies that:

$$\sqrt{\alpha^2 + \frac{\omega^*\alpha}{k}} \left(\sqrt{\alpha^2 + \frac{\omega^*\alpha}{k}} - \alpha \right)^2 = \frac{3B\alpha}{8k} \tag{4}$$

The specific certification process is in Appendix.

The above proposition provides the manufacturer's optimal price decision, and we can obtain the manufacturer's optimal profit under the benchmark case.

4.2 Deadline incentive mechanism

In this case, we consider a situation where the manufacturer offers a deadline incentive mechanism to the suppliers. The manufacturer decides optimal price ω paid to the suppliers, and the suppliers make the optimal work rate decision $\mu_i (i = 1, 2)$. Unlike sub-section 4.1, the two suppliers' decisions are no longer independent of each other under the incentive mechanisms.

Consistent with the above analysis, this incentive mechanism contains the following three components:

A base price. The price manufacturer pays to the supplier is ω , which is irrelevant to the performance of supplier's delivery time.

A specific due date. The manufacturer sets a pre-determined delivery date, which we referred as a deadline D .

An incentive/disincentive scheme. The supplier will be awarded a reward at r per unit time of early completion than the deadline and will be charged a penalty at p per unit time of delay than the deadline.

Under the deadline incentive mechanism, one supplier makes his optimal work rate decision regardless of the other supplier's decision. After the completion of his own subtask, supplier i ($i = 1, 2$) will get an expected payment of $W^d(i)$, which satisfies that:

$$W^d(i) = \omega + rE(D - X_i)^+ - pE(X_i - D)^+ \tag{5}$$

By some calculations, we can get the following:

$$W^d(i) = \omega + pD - pE(X_i) + (r - p) \left[D \int_0^D f(t_i) dt_i - \int_0^D t_i f(t_i) dt_i \right]$$

Lemma 1. The expected payment $W^d(i)$ satisfies that:

$$W^d(i) = \omega + rD - r \frac{1}{\mu_i} + (r - p) \frac{1}{\mu_i} e^{-\mu_i D} \tag{6}$$

See Appendix for the specific certification process.

Furthermore, we can express supplier i 's expected discounted profit as follows:

$$\Pi_i^d = W^d(i)E[e^{-\alpha X_i}] - \frac{k\mu_i^2}{\alpha + \mu_i} \tag{7}$$

By some calculations, we can get the following proposition.

Proposition 3. The suppliers' optimal work rates are the same in equilibrium and it is characterized by:

$$k(\mu^{d*})^2 + 2k\alpha\mu^{d*} + (r - p)e^{-\mu^{d*}D}(1 + D\alpha + D\mu^{d*}) = \omega\alpha + rD\alpha + r \tag{8}$$

The specific certification process is in Appendix.

We define $G = r - p$. From proposition 3, we have that if $G \leq 0$, we can get that $\mu^{d*} > \mu^*$. If $G > 0$, when $e^{-\mu^{d*}D}(1 + D\alpha + D\mu^{d*}) < \frac{rD\alpha + r}{r - p}$, we have $\mu^{d*} > \mu^*$. When $e^{-\mu^{d*}D}(1 + D\alpha + D\mu^{d*}) > \frac{rD\alpha + r}{r - p}$, we have $\mu^{d*} < \mu^*$. In equilibrium, we have $W^d(1) = W^d(2) = W^d$.

We then make some comparative static analyses concerning on the supplier's optimal work rate under the deadline incentive mechanism. Table 2 shows the relationship between supplier's optimal work rate and the pre-determined specific due date.

We then show the relationship between supplier's optimal work rate and the disincentive mechanism factor, and the relationship between supplier's optimal work rate and the base price in Table 3.

As to the relationship between supplier's optimal work rate and the incentive mechanism factor, the results are given in Table 4.

Table 2 The relationship between μ^{d*} and D

$G < 0$	$G = 0$	$G > 0$
$\frac{d}{dD}\mu^{d*} > 0$, if $\mu^{d*} < -\frac{1}{D} \ln \frac{2k}{D^2(r-p)}$	$\frac{d}{dD}\mu^{d*} > 0$	$\frac{d}{dD}\mu^{d*} > 0$, i. if $\mu^{d*} < -\frac{\alpha}{2} + \sqrt{\frac{\alpha^2}{4} + \frac{\alpha}{D}}$
		ii. if $\mu^{d*} > -\frac{\alpha}{2} + \sqrt{\frac{\alpha^2}{4} + \frac{\alpha}{D}}$ and $e^{-\mu^{d*}D}(D\alpha\mu^{d*} + D(\mu^{d*})^2 - \alpha) < \frac{r\alpha}{p-r}$
$\frac{d}{dD}\mu^{d*} < 0$, if $\mu^{d*} > -\frac{1}{D} \ln \frac{2k}{D^2(r-p)}$	$\frac{d}{dD}\mu^{d*} < 0$,	i. if $\mu^{d*} > -\frac{\alpha}{2} + \sqrt{\frac{\alpha^2}{4} + \frac{\alpha}{D}}$ and $e^{-\mu^{d*}D}(D\alpha\mu^{d*} + D(\mu^{d*})^2 - \alpha) > \frac{r\alpha}{p-r}$

Table 3 The relationship between μ^{d*} and p, ω

$G \leq 0$	$G > 0$
$\frac{d}{dp}\mu^{d*} > 0$	$\frac{d}{dp}\mu^{d*} > 0$, if $\mu^{d*} < -\frac{1}{D} \ln \frac{2k}{D^2(r-p)}$ $\frac{d}{dp}\mu^{d*} < 0$, if $\mu^{d*} > -\frac{1}{D} \ln \frac{2k}{D^2(r-p)}$
$\frac{d}{d\omega}\mu^{d*} > 0$	$\frac{d}{d\omega}\mu^{d*} > 0$, if $\mu^{d*} < -\frac{1}{D} \ln \frac{2k}{D^2(r-p)}$ $\frac{d}{d\omega}\mu^{d*} < 0$, if $\mu^{d*} > -\frac{1}{D} \ln \frac{2k}{D^2(r-p)}$

Table 4 The relationship between μ^{d*} and r

$G \leq 0$	$G > 0$
$\frac{d}{dr} \mu^{d*} > 0,$ if $\mu^{d*} < -\alpha + \sqrt{\alpha^2 + \frac{p+D\alpha p + \omega\alpha}{k}}$	$\frac{d}{dr} \mu^{d*} > 0,$ i. if $\mu^{d*} < \min \left\{ -\alpha + \sqrt{\alpha^2 + \frac{p+D\alpha p + \omega\alpha}{k}}, -\frac{1}{D} \ln \frac{2k}{D^2(r-p)} \right\}$ ii. if $\mu^{d*} > \max \left\{ -\alpha + \sqrt{\alpha^2 + \frac{p+D\alpha p + \omega\alpha}{k}}, -\frac{1}{D} \ln \frac{2k}{D^2(r-p)} \right\}$
$\frac{d}{dr} \mu^{d*} < 0,$ if $\mu^{d*} > -\alpha + \sqrt{\alpha^2 + \frac{p+D\alpha p + \omega\alpha}{k}}$	$\frac{d}{dr} \mu^{d*} < 0,$ i. if $\mu^{d*} \sim \left[-\alpha + \sqrt{\alpha^2 + \frac{p+D\alpha p + \omega\alpha}{k}}, -\frac{1}{D} \ln \frac{2k}{D^2(r-p)} \right]$ ii. if $\mu^{d*} \sim \left[-\frac{1}{D} \ln \frac{2k}{D^2(r-p)}, -\alpha + \sqrt{\alpha^2 + \frac{p+D\alpha p + \omega\alpha}{k}} \right]$

After the delivery of subtasks by the two suppliers, the outsourcing activities ends. The manufacturer’s profit can be expressed as

$$\Pi_m^d = A - BE[T] - 2W^d. \tag{9}$$

In order to solve the above problem, we transfer (8) as $L(\mu^{d*}, \omega) = 0$. Manufacturer’s profit maximizing problem in Eq. 9 can be redeemed as an optimization problem, which can be recognized as follows:

$$\begin{aligned} \max_{\mu^{d*}, \omega} \Pi_m^d &= A - BE[T] - 2W^d \\ \text{s.t. } L(\mu^{d*}, \omega) &= 0 \end{aligned}$$

Because of the implicit function form of Eq. 8, it’s hard to get the analytical expression of optimal price decision made by the manufacturer. We do some numerical examples in Section 5 to determine if deadline incentive mechanisms dominate the others. We use optimization algorithms to determine the optimal (μ^{d*}, ω) . We do some comparative analyses concerning supplier’s optimal work rate decision in this sub-section and discuss the other two incentive mechanisms in the following analysis.

4.3 Competition mechanism

In this section, we consider a situation where the manufacturer offers a competition mechanism to the suppliers. The manufacturer makes a decision of optimal price ω paid to the suppliers, and the suppliers make the optimal work rate decision $\mu_i (i = 1,2)$. The two suppliers also have to take the other’s decision into his own consideration under this incentive mechanisms.

This incentive mechanism manufacturer offers to the suppliers contains the following three components:

A base price. The manufacturer pays each supplier a base price of ω .

A comparison of the faster and slower. Because of the stochasticity of project duration times, the result ex post can be distinguished from a faster supplier and a slower supplier.

An incentive/disincentive scheme. The faster supplier will be awarded a reward of r per unit time earlier than the slower supplier; the slower supplier will be charged a penalty of p per unit time later than the faster supplier.

For supplier $i (i = 1,2)$, we can get that his expected payment gained from the manufacturer after his completion of the subtask satisfies:

$$W^c(i) = \omega + rE[X_{-i} - X_i]^+ - pE[X_i - X_{-i}]^+ \tag{10}$$

Lemma 2. With simple calculations, we can infer that:

$$W^c(i) = \omega + r \frac{\mu_i}{\mu_{-i}(\mu_i + \mu_{-i})} - p \frac{\mu_{-i}}{\mu_i(\mu_i + \mu_{-i})}$$

The specific certification process is in Appendix.

Furthermore, we have that supplier i ’s expected discounted profit can be expressed as:

$$\Pi_i^c = W^c(i)E[e^{-\alpha X_i}] - \frac{k\mu_i^2}{\alpha + \mu_i} \tag{11}$$

Proposition 4. In equilibrium, supplier’s optimal work rate is given by:

$$4k(\mu^{c*})^4 + 8k\alpha(\mu^{c*})^3 - (3p + 4\omega\alpha)(\mu^{c*})^2 + (r - p\alpha)\mu^{c*} = r\alpha \tag{12}$$

See Appendix for the specific certification process.

From the above result, we can see that if $p = r = 0$, we have $\mu^{c*} = \mu^*$. Furthermore, we have the following:

If $\frac{7-4\sqrt{3}}{\alpha} < \frac{p}{r} < \frac{7+4\sqrt{3}}{\alpha}$, we have $\mu^{c*} > \mu^*$.

If $\frac{p}{r} < \frac{7-4\sqrt{3}}{\alpha}$ or $\frac{p}{r} > \frac{7+4\sqrt{3}}{\alpha}$, we have when $\mu^{c*} < \frac{r-p\alpha-\sqrt{p^2\alpha^2+r^2-14par}}{6p}$ or $\mu^{c*} > \frac{r-p\alpha+\sqrt{p^2\alpha^2+r^2-14par}}{6p}$, then $\mu^{c*} > \mu^*$; when $\frac{r-p\alpha-\sqrt{p^2\alpha^2+r^2-14par}}{6p} < \mu^{c*} < \frac{r-p\alpha+\sqrt{p^2\alpha^2+r^2-14par}}{6p}$, then $\mu^{c*} < \mu^*$. And in equilibrium, we have $W^c(1) = W^c(2) = W^c$.

We then make some comparative static analyses under the competitive mechanism, and show the relationship between supplier’s optimal work rate and the base price, the incentive factor and the disincentive factor in Table 5.

Table 5 comparative static analyses under competition model

$\frac{d}{dr}\mu^{c*}$	$\frac{d}{dr}\mu^{c*} > 0$,
	i. if $\mu^{c*} < \alpha$ and $16k(\mu^{c*})^3 + 24kr(\mu^{c*})^2 - 8\omega\alpha\mu^{c*} - 6p\mu^{c*} > p\alpha - r$ ii. if $\mu^{c*} > \alpha$ and $16k(\mu^{c*})^3 + 24kr(\mu^{c*})^2 - 8\omega\alpha\mu^{c*} - 6p\mu^{c*} < p\alpha - r$
$\frac{d}{dp}\mu^{c*}$	$\frac{d}{dp}\mu^{c*} > 0$, if $16k(\mu^{c*})^3 + 24kr(\mu^{c*})^2 - 8\omega\alpha\mu^{c*} - 6p\mu^{c*} > p\alpha - r$ $\frac{d}{dp}\mu^{c*} < 0$, if $16k(\mu^{c*})^3 + 24kr(\mu^{c*})^2 - 8\omega\alpha\mu^{c*} - 6p\mu^{c*} < p\alpha - r$
	$\frac{d}{d\omega}\mu^{c*} > 0$, if $16k(\mu^{c*})^3 + 24kr(\mu^{c*})^2 - 8\omega\alpha\mu^{c*} - 6p\mu^{c*} > p\alpha - r$ $\frac{d}{d\omega}\mu^{c*} < 0$, if $16k(\mu^{c*})^3 + 24kr(\mu^{c*})^2 - 8\omega\alpha\mu^{c*} - 6p\mu^{c*} < p\alpha - r$

After both subtasks are completed, the outsourcing activity ends. We can express the manufacturer’s profit as follows:

$$\Pi_m^c = A - BE[T] - 2W^c \tag{13}$$

We redefine Eq. 12 as $L(\mu^{c*}, \omega) = 0$. Furthermore, the manufacturer’s problem can be re-deemed as:

$$\begin{aligned} \max_{\mu^{c*}, \omega} \Pi_m^c &= A - BE[T] - 2W^c \\ \text{s.t. } L(\mu^{c*}, \omega) &= 0 \end{aligned}$$

Due to the complexity of Eq. 12, the analytical expression of the optimal price decision made by the manufacturer is hard to determine. Some numerical expression is provided in Section 5. We seek to compare the results among different models with respect to manufacturer’s profit function.

4.4 Mixed mechanism

In this case, we consider an incentive mechanism combining the above analyzed deadline and competition incentives. The manufacturer makes a decision of optimal price ω paid to the suppliers, and the suppliers make the optimal work rate decision $\mu_i (i = 1, 2)$.

The mixed incentive mechanism manufacturer can offer to the suppliers contains the following four components:

A base price. The manufacturer pays a base price of ω to the suppliers.

A specific due date. The manufacturer sets a pre-determined delivery due date D .

A comparison of the faster and slower. Due to the randomness of project completion times, the result ex post can be distinguished from a faster supplier and a slower supplier.

An incentive/disincentive scheme. The faster supplier will be awarded a reward of r per unit time earlier than the deadline; the slower supplier will be charged a penalty of p per unit time later than the deadline.

For supplier i ($i = 1, 2$), after his completion of his own subtasks, he receives the pre-agreed base price ω . We can get that his expected payment gained from the manufacturer satisfies:

$$W^m(i) = \omega + rE[(D - X_i)^+ | X_i < X_{-i}] - pE[(X_i - D)^+ | X_i > X_{-i}] \tag{14}$$

Lemma 3. With some algebra calculations, we can get the following:

$$W^m(i) = \omega + rD - r \frac{1}{\mu_i + \mu_{-i}} + r \frac{1}{\mu_i + \mu_{-i}} e^{-(\mu_i + \mu_{-i})D} - p \frac{1}{\mu_{-i}} \left[\frac{\mu_i + \mu_{-i}}{\mu_i} e^{-\mu_i D} - \frac{\mu_i}{\mu_i + \mu_{-i}} e^{-(\mu_i + \mu_{-i})D} \right]$$

See Appendix for the certification process.

Furthermore, supplier i 's expected discounted profit can be expressed as follows:

$$\Pi_i^m = W^m(i)E[e^{-\alpha X_i}] - \frac{k\mu_i^2}{\alpha + \mu_i} \tag{15}$$

Proposition 5. The first-order condition of supplier's profit satisfies that:

$$4k(\mu^{m*})^4 + \Phi(\mu^{m*})^3 - \Psi(\mu^{m*})^2 - \Upsilon\mu^{m*} = r\alpha - r\alpha e^{-2\mu^{m*}D} \tag{16}$$

in which $\Phi = 8k\alpha - 8pDe^{-\mu^{m*}D} + 2pDe^{-2\mu^{m*}D}$, $\Psi = 4\omega\alpha + 4rD\alpha - 2rDe^{-2\mu^{m*}D} + 8p\alpha De^{-\mu^{m*}D} - 2p\alpha De^{-2\mu^{m*}D} + 4pe^{-\mu^{m*}D} + pe^{-2\mu^{m*}D}$, and $\Upsilon = 3r - 3re^{-2\mu^{m*}D} - 2rD\alpha e^{-2\mu^{m*}D} - 4p\alpha e^{-\mu^{m*}D} + 3p\alpha e^{-2\mu^{m*}D}$.

The specific certification process is in Appendix.

If $r = p = 0$, we have that $\mu^{m*} = \mu^*$. In equilibrium, we have $W^m(1) = W^m(2) = W^m$. We can infer that if formula Eq. 15 has a maximum value, then supplier's optimal work rate satisfies Eq. 16; otherwise, we can find the maximum value of Eq. 15 by limiting the scope of parameter ω .

Furthermore, we have that the manufacturer's profit satisfies that:

$$\Pi_m^m = A - BE[T] - 2W^m \tag{17}$$

In the next section, we also provide the manufacturer's optimal price decision through some numerical examples. Section 4 provides three kinds of incentive mechanisms to compare with the benchmark case. Deadline, competition, and mixed incentive mechanisms are analyzed and the comparison results are given in the following analysis.

5. Numerical analysis

Due to the complexity of the manufacturer's profit function, we cannot get the analytical expression of manufacturer's optimal price decision under the three proposed incentive mechanism models. We compare the results among different models through some numerical examples.

In our numerical examples, we set $\alpha = 1, k = 1, A = 40$ and $B = 1$. We use some abbreviations to represent different incentive models, i.e., BM is short for base model, DM is short for deadline incentive model, CM is short for competition model and MM short for mixed model.

Fig. 1 shows the results among different models when parameter p varies from 0 to 10 given parameter $D = 0.5$. With different values of parameter r , i.e., $r = 0, 2, 4, 6, 8, 10$, we get six numerical examples, and manufacture's optimal profit among different incentive mechanisms differs under different circumstances.

Observation 1. The comparison results among different incentive models differ concerning different absolute value and relative relationship between parameter r and p . All four mechanisms may be in the domination position under different circumstances.

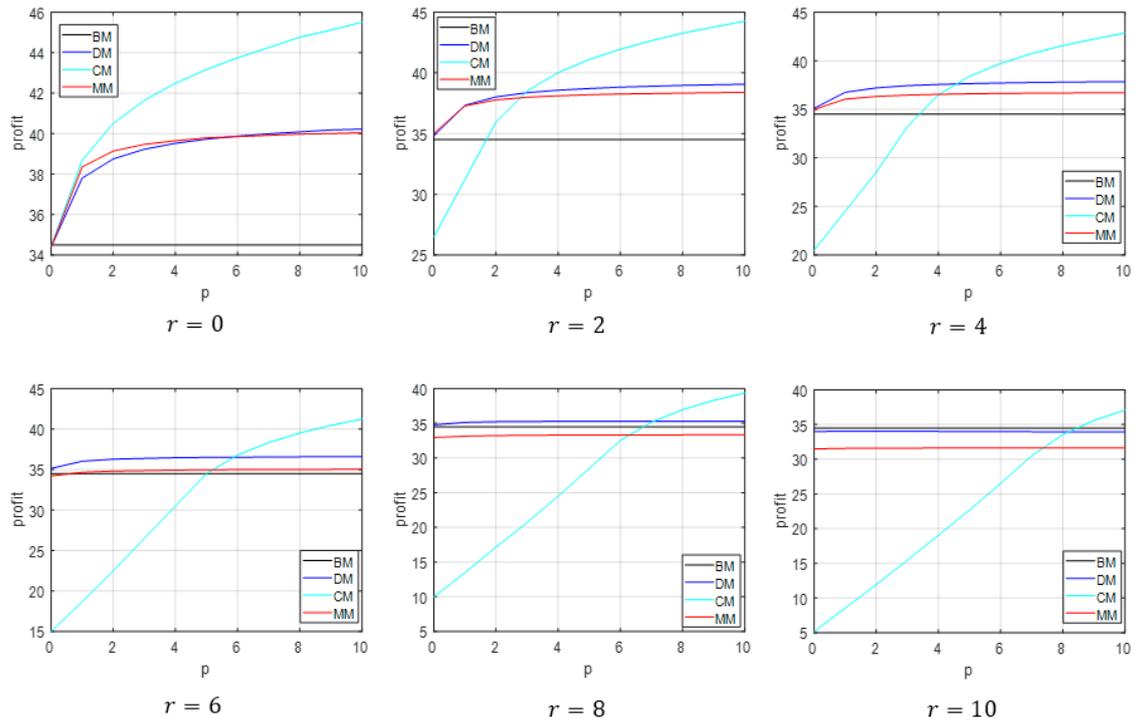


Fig. 1 Numerical examples with varying parameter r with given $D = 0.5$

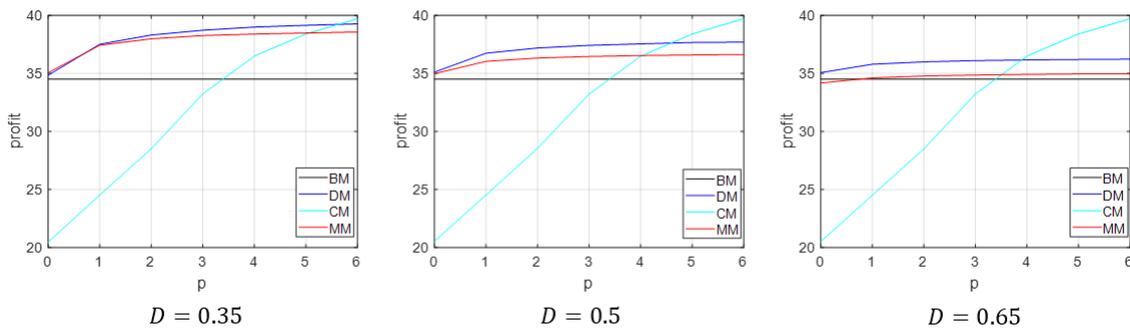


Fig. 2 Numerical examples with varying parameter D with given $r = 4$

We then show the comparison results among different models with a different specific deadline. Fig. 2 gives the specific results with given parameter $r = 4$, and p varies from 0 to 10. By choosing parameter D among 0.35, 0.5 and 0.65, we get three numerical examples.

Observation 2. Given the absolute and relative value of parameters r and p , the influence of specific due date D has a different impact on the comparison results among different incentive models. A smaller D has a same impact on the results with a smaller r when other parameters are the same; a larger D has a same impact on the results with a larger r when other parameters are the same.

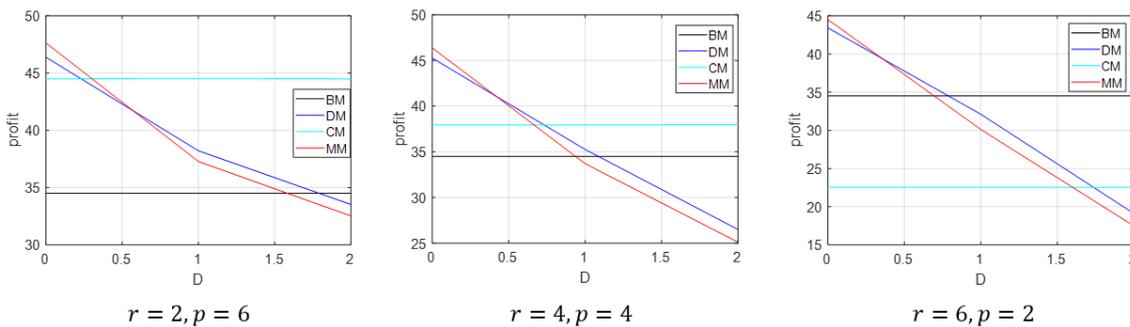


Fig. 3 Numerical examples with varying parameter r and p

Fig. 3 shows the results with varying parameter D from 0 to 2. With 3 different pairs of parameters r and p , we get different numerical results on manufacturer's optimal choice from choosing different incentive mechanisms.

Observation 3. Different combination pairs of parameters r and p has a bigger influence on the performance of incentive model CM. The performance of incentive models DM and MM decreases with the specific due date D . Model MM dominates DM when parameter D is smaller and model DM nominates MM when parameter D is larger.

In this Section, we show the comparative results through some numerical examples, and the meaning results can shed light on some managerial implications for the real-world business practice.

6. Conclusion

We consider an incentive mechanism design problem faced by a manufacturer when conducting outsourcing activities. Three different kinds of incentive mechanisms are proposed in our article. We get some comparative results and management implications for the real-world business practice.

6.1 Findings

We proposed three kinds of incentive mechanisms. The first one is deadline incentive mechanism, in which the manufacturer impose a specific deadline and when the supplier completes earlier than the deadline will be rewarded and will be penalized when completes later than the deadline.

The second incentive mechanism is competition mechanism. In our setting where the manufacturer outsources two parallel subtasks to two different suppliers, the two suppliers may end the corresponding subtask with different duration times due to the stochasticity of the project.

The third incentive mechanism combines the above two mechanisms, which we referred to as mixed incentive mechanism. It considers both the deadline in projects and competitiveness in parallel projects. Only the faster supplier is awarded when earlier than the deadline and only the slower supplier is penalized when later than the deadline.

We find that the results of the comparison of the three mechanisms depend on the related parameters settings, and are all likely to be in a dominant position. We give the specific results through some numerical examples.

We assume that the supplier can get a base price ω in the above analysis, and different base price the manufacturer paid to the supplier is related to different project size. Thus, our findings can be generalized to projects of different magnitudes, which is meaningful to the real-world business practice.

6.2 Future research

There are some extensions that we can continue to work on in the future. We list some possible research schemes in the following content.

In our work, we consider a two parallel sub-projects structure, however, what happens when there are more than two tasks in parallel? It could be a lot more complicated and needs further study by scholars. We can figure that the deadline incentive mechanism can gain the same result as we show in sub-section 4.2. Regarding to competition incentive mechanism, different from two parallel sub-projects structure, when there are n ($n \geq 2$) tasks in parallel, we can set that the fastest supplier can get a reward and the slowest supplier will get a penalty. The specific incentive result needs our further study. As to the mixed incentive mechanism, we can also consider that the fastest supplier can get a reward if his completion time is earlier than the deadline and vice versa. However, in this case, the modeling process may be very complicated and whether we can gain some insightful results is unclear.

Serial project structure is another direction we can pay attention to. Also, we are wondering that what happens when there are more than two tasks in a general assembly network? Do any

of the results reported in this paper continue to hold? When involving network structure, the study will be more complicated and difficult.

The above discussion prompts us to pay more attention to the problems in project supply chain management and try to adopt new methods to solve these problems. In addition, emerging applications in project management will certainly identify further important research questions and opportunities in addition to the issues we discuss here. With the increasing complexity of project supply chain and the increasing subcontracting activities, how to solve this problem is a great challenge. Therefore, it requires the attention and time of our scholars to conduct research and demonstrate some management implications to guide guidance into the real world.

We believe that in the future, scholars will make great progress in academic research in this field, which will further guide the practical business application of project management. In the next several years we will have more and more important results, which are of course the basis for future research.

Acknowledgement

This work was supported by the Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJQN202200808).

References

- [1] Boarnet, M.G. (1998). Business losses, transportation damage and the Northridge earthquake, *Journal of Transportation and Statistics*, Vol. 1, No. 2, 49-64, [doi: 10.21949/1501575](https://doi.org/10.21949/1501575).
- [2] Bayiz, M., Corbett, C.J. (2005). Coordination and incentive contracts in project management under asymmetric information, Working paper, UCLA Anderson School, [doi: 10.2139/ssrn.914227](https://doi.org/10.2139/ssrn.914227).
- [3] Lindkvist, L., Soderlund, J., Tell, F. (1998). Managing product development projects: On the significance of fountains and deadlines, *Organization Studies*, Vol. 19, No. 6, 931-951, [doi: 10.1177/017084069801900602](https://doi.org/10.1177/017084069801900602).
- [4] Kerzner, H. (2017). *Project management: A systems approach to planning, scheduling, and controlling*, John Wiley & Sons, Hoboken, New Jersey, USA.
- [5] Eisenhardt, K.M., Tabrizi, B.N. (1995). Accelerating adaptive processes: Product innovation in the global computer industry, *Administrative Science Quarterly*, Vol. 40, No. 1, 84-110, [doi: 10.2307/2393701](https://doi.org/10.2307/2393701).
- [6] Su, J., Zhang, F., Wang, D., Sindakis, S., Xiao, Y., Herrera-Viedma, E. (2023). Examining the influence of knowledge spillover on partner selection in knowledge Alliances: The role of benefit distribution, *Computers & Industrial Engineering*, Vol. 180, Article No. 109245, [doi: 10.1016/j.cie.2023.109245](https://doi.org/10.1016/j.cie.2023.109245).
- [7] Nahmias, S., Olsen, T.L. (2015). *Production and operations analysis*, Seventh Edition, Waveland Press, Long Grove, Illinois, USA.
- [8] Hall, N.G. (2016). Research and teaching opportunities in project management, *INFORMS TutORials in Operations Research*, 329-388, [doi: 10.1287/educ.2016.0146](https://doi.org/10.1287/educ.2016.0146).
- [9] Zhang, J. (2016). Deadlines in product development, *Management Science*, Vol. 62, No. 11, 3310-3326, [doi: 10.1287/mnsc.2015.2300](https://doi.org/10.1287/mnsc.2015.2300).
- [10] Bordley, R.F., Keisler, J.M., Logan, T.M. (2019). Managing projects with uncertain deadlines, *European Journal of Operational Research*, Vol. 274, No. 1, 291-302, [doi: 10.1016/j.ejor.2018.09.036](https://doi.org/10.1016/j.ejor.2018.09.036).
- [11] Du, L., Hu, M., Wu, J. (2022). Contingent stimulus in crowdfunding, *Production & Operations Management*, Vol. 31, No. 9, 3543-3559, [doi: 10.1111/poms.13782](https://doi.org/10.1111/poms.13782).
- [12] Gupta, D., Snir, E.M., Chen, Y. (2015). Contractors' and agency decisions and policy implications in A+B bidding, *Production and Operations Management*, Vol. 24, No. 1, 159-177, [doi: 10.1111/poms.12217](https://doi.org/10.1111/poms.12217).
- [13] Tang, C.S., Zhang, K., Zhou, S.X. (2015). Incentive contracts for managing a project with uncertain completion time, *Production and Operations Management*, Vol. 24, No. 12, 1945-1954, [doi: 10.1111/poms.12387](https://doi.org/10.1111/poms.12387).
- [14] Chen, S., Lee, H. (2016). Incentive alignment and coordination of project supply chains, *Management Science*, Vol. 63, No. 4, 1011-1025, [doi: 10.1287/mnsc.2015.2373](https://doi.org/10.1287/mnsc.2015.2373).
- [15] Kwon, H.D., Lippman, S.A., Tang, C.S. (2010). Optimal time-based and cost-based coordinated project contracts with unobservable work rates, *International Journal of Production Economics*, Vol. 126, No. 2, 247-254, [doi: 10.1016/j.ijpe.2010.03.013](https://doi.org/10.1016/j.ijpe.2010.03.013).
- [16] Zhao, J., Mu, Y. (2022). Managing projects under different payment schemes, In: *Proceedings of the 5th European International Conference on Industrial Engineering and Operations Management*, Rome, Italy, [doi: 10.46254/EU05.20220282](https://doi.org/10.46254/EU05.20220282).
- [17] Bergantiños, G., Lorenzo, L. (2019). How to apply penalties to avoid delays in projects, *European Journal of Operational Research*, Vol. 275, No. 2, 608-620, [doi: 10.1016/j.ejor.2018.11.056](https://doi.org/10.1016/j.ejor.2018.11.056).
- [18] Estévez-Fernández, A. (2012). A game theoretical approach to sharing penalties and rewards in projects, *European Journal of Operational Research*, Vol. 216, No. 3, 647-657, [doi: 10.1016/j.ejor.2011.08.015](https://doi.org/10.1016/j.ejor.2011.08.015).

- [19] Chen, T., Klastorin, T., Wagner, M.R. (2015). Incentive contracts in serial stochastic projects, *Manufacturing & Service Operations Management*, Vol. 17, No. 3, 290-301, doi: 10.1287/msom.2015.0528.
- [20] Kwon, H.D., Lippman, S.A., McCardle, K., Tang, C.S. (2010). Managing project contracts with delayed payments, *Manufacturing & Service Operations Management*, Vol. 12, No. 4, 692-707, doi: 10.1287/msom.1100.0301.
- [21] Song, J.M., Zhao, Y., Xu, X. (2020). Incentives and gaming in collaborative projects under risk-sharing partnerships, *Manufacturing & Service Operations Management*, Vol. 23, No. 2, 453-470, doi: 10.1287/msom.2019.0840.
- [22] Dawande, M., Janakiraman, G., Qi, A., Wu, Q. (2019). Optimal incentive contracts in project management, *Production and Operations Management*, Vol. 28, No. 6, 1431-1445, doi: 10.1111/poms.12997.

Appendix

Proof of Proposition 2. By taking the first-order condition of manufacturer’s profit function, we have:

$$\frac{d}{d\omega} \Pi_m = -2 + \frac{3B}{2\left(-\alpha + \sqrt{\alpha^2 + \frac{\omega\alpha}{k}}\right)^2} \cdot \frac{\alpha}{2k\sqrt{\alpha^2 + \frac{\omega\alpha}{k}}}$$

By taking the second-order condition of Π_m , we can obtain:

$$\frac{d^2}{d\omega^2} \Pi_m = \frac{3B\alpha}{4k} \cdot \frac{-1}{2\left(-\alpha + \sqrt{\alpha^2 + \frac{\omega\alpha}{k}}\right)^3} \cdot \left(\frac{\alpha}{2c\sqrt{\alpha^2 + \frac{\omega\alpha}{k}}}\right)^2 + \frac{3B\alpha}{4k} \cdot \frac{1}{\left(-\alpha + \sqrt{\alpha^2 + \frac{\omega\alpha}{k}}\right)^2} \cdot \frac{-1}{2\left(\alpha^2 + \frac{\omega\alpha}{k}\right)^{\frac{3}{2}}}$$

We can get that $\frac{d^2}{d\omega^2} \Pi_m < 0$. Thus, we can infer that Π_m is a concave function with respect to parameter ω .

Proof of Lemma 1. From $W^d = \omega + rE(D - X_i)^+ - pE(X_i - D)^+$, we can get:

$$W^d = \omega + r \int_0^D (D - t_i) f(t_i) dt_i - p \int_D^\infty (t_i - D) f(t_i) dt_i = \omega + pD - pE(X_i) + (r - p) \left[D \int_0^D f(t_i) dt_i - \int_0^D t_i f(t_i) dt_i \right].$$

Thus, we have $W^d = +rD - r \frac{1}{\mu_i} + (r - p) \frac{1}{\mu_i} e^{-\mu_i D}$, and the lemma is proved.

Proof of Proposition 3. By some calculations, we can obtain:

$$\begin{aligned} \frac{d}{d\mu_i} \Pi_i^d &= \frac{1}{(\alpha + \mu_i)^2} [\omega\alpha + rD\alpha + r - (r - p)e^{-\mu_i D} (1 + D\alpha + D\mu_i) - k(\mu_i)^2 - 2k\alpha\mu_i]. \\ \frac{d^2}{d(\mu_i)^2} \Pi_i^d &= \frac{1}{(\alpha + \mu_i)^3} [(r - p)e^{-\mu_i D} (D\alpha + D\mu_i) D(\alpha + \mu_i) + 2(r - p)e^{-\mu_i D} (1 + D\alpha + D\mu_i) - 2c\alpha^2 - 2(\omega\alpha + rD\alpha + r)]. \end{aligned}$$

If $r \leq p$, we can get that Π_i^d is a concave function. Thus, we can get the result by taking the first order condition of (7).

If $r > p$, for sufficiently large μ_i , we have $\frac{d}{d\mu_i} \Pi_i^d < 0$. And we can infer that $\lim_{\mu_i \rightarrow 0} \frac{d}{d\mu_i} \Pi_i^d = \frac{1}{\alpha^2} [\omega\alpha + rD\alpha + r - (r - p)(1 + D\alpha)] = \omega\alpha + p + pD\alpha > 0$.

Therefore, we can infer that there exists a maximum value of the supplier’s profit function and the proposition is proved.

Proof of Lemma 2. Let $Z = X_i - X_{-i}$, thus, the probability density function of random variable Z when $Z > 0$ satisfies that

$$f(z) = \int_{t_i=z}^\infty \mu_i e^{-\mu_i t_i} \mu_{-i} e^{-\mu_{-i}(t_i-z)} dt_i = \frac{\mu_i \mu_{-i}}{(\mu_i + \mu_{-i})} e^{-\mu_i z}.$$

Therefore, we can get that

$$E[X_i - X_{-i}]^+ = \int_0^\infty z f(z) dz = \frac{\mu_{-i}}{\mu_i(\mu_i + \mu_{-i})}.$$

Thus, the lemma is proved.

Proof of Proposition 4. From formula Eq. 12, we have

$$\frac{d}{d\mu_i} \Pi_i^c = \frac{1}{(\alpha + \mu_i)^2 (\mu_i + \mu_{-i})^2} \left[\omega\alpha(\mu_i + \mu_{-i})^2 + r(\alpha + \mu_i) - r \frac{\mu_i}{\mu_{-i}} (\mu_i + \mu_{-i}) + p\mu_{-i}(\alpha + 2\mu_i + \mu_{-i}) - (k\mu_i^2 + 2k\alpha\mu_i)(\mu_i + \mu_{-i})^2 \right].$$

For sufficiently large μ_i , we have $\frac{d}{d\mu_i} \Pi_i^c < 0$. And we can get that $\lim_{\mu_i \rightarrow 0} \frac{d}{d\mu_i} \Pi_i^c = \frac{1}{\alpha^2 \mu_{-i}^2} [\omega \alpha \mu_{-i}^2 + r \alpha + p \mu_{-i} (\alpha + \mu_{-i})] > 0$.

Thus, we can infer that there exists a maximum value of the supplier's profit function and the proposition is proved.

Proof of Lemma 3. The conditional distribution function $F(X_i < t | X_i < X_{-i}) = \frac{\text{Pro}(X_i < t, X_i < X_{-i})}{\text{Pro}(X_i < X_{-i})}$.

We first examine that $\text{Pro}(X_i < t, X_i < X_{-i}) = \int_0^t \mu_i e^{-\mu_i t_i} \int_{t_i}^\infty \mu_{-i} e^{-\mu_{-i} t_{-i}} dt_{-i} dt_i = \frac{\mu_i}{\mu_i + \mu_{-i}} [1 - e^{-(\mu_i + \mu_{-i})t}]$.

By calculation, we can infer that $\text{Pro}(X_i < X_{-i}) = \int_0^\infty \mu_{-i} e^{-\mu_{-i} t_{-i}} \int_0^{t_{-i}} \mu_i e^{-\mu_i t_i} dt_i dt_{-i} = \frac{\mu_i}{\mu_i + \mu_{-i}}$.

Thus, we have $F(X_i < t | X_i < X_{-i}) = 1 - e^{-(\mu_i + \mu_{-i})t}$.

Thus, we can get that $E[(D - X_i)^+ | X_i < X_{-i}] = (\mu_i + \mu_{-i}) \int_0^D (D - t) e^{-(\mu_i + \mu_{-i})t} dt = D + \frac{1}{\mu_i + \mu_{-i}} [e^{-(\mu_i + \mu_{-i})D} - 1]$.

By the same way, we can get that $F(X_i < t | X_i > X_{-i}) = \frac{\text{Pro}(X_i < t, X_i > X_{-i})}{\text{Pro}(X_i > X_{-i})}$. Furthermore, we have

that $\text{Pro}(X_i < t, X_i > X_{-i}) = \int_0^t \mu_i e^{-\mu_i t_i} \int_0^{t_i} \mu_{-i} e^{-\mu_{-i} t_{-i}} dt_{-i} dt_i = \frac{\mu_{-i}}{\mu_i + \mu_{-i}} - e^{-\mu_i t} + \frac{\mu_i}{\mu_i + \mu_{-i}} e^{-(\mu_i + \mu_{-i})t}$, and we can infer that $\text{Pro}(X_i > X_{-i}) = \frac{\mu_{-i}}{\mu_i + \mu_{-i}}$.

Thus, we can get that $E[(X_i - D)^+ | X_i > X_{-i}] = \frac{\mu_i}{\mu_{-i}} (\mu_i + \mu_{-i}) \int_D^\infty (t - D) [e^{-\mu_i t} - e^{-(\mu_i + \mu_{-i})t}] dt = \frac{1}{\mu_{-i}} [\frac{\mu_i + \mu_{-i}}{\mu_i} e^{-\mu_i D} - \frac{\mu_i}{\mu_i + \mu_{-i}} e^{-(\mu_i + \mu_{-i})D}]$.

Thus, the lemma is proved.

Proof of Proposition 5. From the supplier's profit function, we can get

$$\begin{aligned} \frac{d}{d\mu_i} \Pi_i^m &= \frac{1}{(\alpha + \mu_i)^2 (\mu_i + \mu_{-i})^2} \left\{ (\omega \alpha + r D \alpha) (\mu_i + \mu_{-i})^2 - r (\alpha + \mu_i) (e^{-(\mu_i + \mu_{-i})D} - 1) - r (\mu_i + \mu_{-i}) (e^{-(\mu_i + \mu_{-i})D} - 1) - r D (\alpha + \mu_i) (\mu_i + \mu_{-i}) e^{-(\mu_i + \mu_{-i})D} - p \frac{\alpha}{\mu_{-i}} (\mu_i + \mu_{-i})^2 \left(\frac{\mu_i + \mu_{-i}}{\mu_i} e^{-\mu_i D} - \frac{\mu_i}{\mu_i + \mu_{-i}} e^{-(\mu_i + \mu_{-i})D} \right) - p \frac{\mu_i}{\mu_{-i}} (\alpha + \mu_i) (\mu_i + \mu_{-i})^2 \left[-D \frac{\mu_i + \mu_{-i}}{\mu_i} e^{-\mu_i D} - \frac{\mu_{-i}}{\mu_i^2} e^{-\mu_i D} + D \frac{\mu_i}{\mu_i + \mu_{-i}} e^{-(\mu_i + \mu_{-i})D} - \frac{\mu_{-i}}{(\mu_i + \mu_{-i})^2} e^{-(\mu_i + \mu_{-i})D} \right] - (k \mu_i^2 + 2k \alpha \mu_i) (\mu_i + \mu_{-i})^2 \right\}. \end{aligned}$$

By letting $\mu_i = \mu_{-i} = \mu^{m*}$, we can get formula Eq. 16.

Thus, the proposition is proved.

Reduction of surface defects by optimization of casting speed using genetic programming: An industrial case study

Kovacic, M.^{a,b,c,*}, Zuperl, U.^d, Gusel, L.^d, Brezocnik, M.^d

^aŠTORE STEEL, d.o.o., Research and Development, Štore, Slovenia

^bUniversity of Ljubljana, Faculty of Mechanical Engineering, Laboratory for Fluid Dynamics and Thermodynamics, Ljubljana, Slovenia

^cCollege of Industrial Engineering, Celje, Slovenia

^dUniversity of Maribor, Faculty of mechanical engineering, Maribor, Slovenia

ABSTRACT

Štore Steel Ltd. produces more than 200 different types of steel with a continuous caster installed in 2016. Several defects, mostly related to thermomechanical behaviour in the mould, originate from the continuous casting process. The same casting speed of 1.6 m/min was used for all steel grades. In May 2023, a project was launched to adjust the casting speed according to the casting temperature. This adjustment included the steel grades with the highest number of surface defects and different carbon content: 16MnCrS5, C22, 30MnVS5, and 46MnVS5. For every 10 °C deviation from the prescribed casting temperature, the speed was changed by 0.02 m/min. During the 2-month period, the ratio of rolled bars with detected surface defects (inspected by an automatic control line) decreased for the mentioned steel grades. The decreases were from 11.27 % to 7.93 %, from 12.73 % to 4.11 %, from 16.28 % to 13.40 %, and from 25.52 % to 16.99 % for 16MnCrS5, C22, 30MnVS5, and 46MnVS5, respectively. Based on the collected chemical composition and casting parameters from these two months, models were obtained using linear regression and genetic programming. These models predict the ratio of rolled bars with detected surface defects and the length of detected surface defects. According to the modelling results, the ratio of rolled bars with detected surface defects and the length of detected surface defects could be minimally reduced by 14 % and 189 %, respectively, using casting speed adjustments. A similar result was achieved from July to November 2023 by adjusting the casting speed for the other 27 types of steel. The same was predicted with the already obtained models. Genetic programming outperformed linear regression.

ARTICLE INFO

Keywords:

Continuous casting of steel;
Surface defects;
Automatic control;
Machine learning;
Modelling;
Optimization;
Prediction;
Linear regression;
Genetic programming

*Corresponding author:

miha.kovacic@store-steel.si
(Kovacic, M.)

Article history:

Received 3 November 2023

Revised 15 December 2023

Accepted 21 December 2023



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

1. Introduction

Modern steel production is hard to imagine without continuous casting. However, due to the thermomechanical behaviour during continuous casting, especially in the mould, several types of defects can occur on the cast material. These defects can also manifest on the rolled material. They can be reduced or eliminated with several approaches:

- Optimization of casting equipment (e.g., tundish, submerged entry nozzles, mould or water sprays geometry, casting powder, tundish powder),
- Optimization of secondary metallurgy (e.g., deoxidation, refinement, homogenization, stirring), and

- Optimization of casting parameters (e.g., casting temperature, speed, cooling water flow and pressure).

In the literature, adjustments to the casting speed have been discussed in relation to casting defects (e.g., surface defects [1-13], breakouts [9, 14, 15]) and productivity [1, 3, 14-16]. These discussions have involved individual steel grades [1, 9] or have been more general [3-5, 10-12, 15-17]. Additional factors such as the melt level in the mould [17, 18], melt level fluctuations [12, 17, 18], and melt flow [3] have also been examined. Artificial intelligence approaches have been utilized as well [4, 7, 10, 16, 17]. Unfortunately, these attempts have only been practically implemented in an industrial environment a few times [11, 15, 17].

This article investigates the reduction of surface defects, detected during the examination (via an automatic control line) of the rolled material, using casting speed adjustments based on the casting temperature. The average casting speed adjustments were calculated for several different grades using an in-house developed solidification model [19, 20]. The primary objective during simulations was to maintain the same metallurgical length (i.e., the distance from the mould to the location where the entire melt solidifies) by adjusting the casting speed when the casting temperature changed. Based on several simulations, for every 10 °C deviation from the prescribed casting temperature, the speed was adjusted by 0.02 m/min. The casting temperature was determined based on the content of carbon, sulphur, and aluminium, the liquidus temperature (calculated using the Wensel equation [21]), and the number of cast sequences (several batches – individual melts are continuously cast without interruption).

The materials and methods section discusses the significance of data collection. Subsequently, this data is utilized to predict the ratio of rolled bars with detected surface defects and their lengths using linear regression and genetic programming methods. The results are then analysed and implemented in practice. Finally, the general results are analysed and conclusions are drawn.

2. Materials and methods

Production at the Štore Steel plant begins with scrap melting in an electric arc furnace, followed by tapping, ladle treatment (i.e., secondary metallurgy), and continuous casting of billets measuring 180 mm × 180 mm. The billets can undergo additional heat treatment or be cooled under hoods. Before rolling, they are reheated. The rolled bars can then be straightened, examined for inner soundness and surface quality, cut, sawn, chamfered, drilled, and peeled.

Since March 2016, a new two-strand continuous casting machine with a radius of 9 m has been in operation. The solidification process involves a water-cooled copper mould for primary cooling, water sprays for secondary cooling, and air cooling for tertiary cooling.

During primary cooling, the solidified shell is subjected to thermomechanical stresses, which can lead to numerous casting defects. In the case of square billets, non-uniform shell solidification (i.e., non-uniform heat removal in the mould) can result in rhombic distortion (i.e., rhomboidity). This distortion can cause off-corner cracks, which manifest as longitudinally opened surface defects on the rolled material. These defects are typically detected during an automatic control line examination.

The origin of open surface defects on rolled material can be confirmed based on metallurgical reports, which include analyses of billets and both flat and round bars, dating back to the initiation of the continuous caster in 2016. Fig. 1 displays a macro-etched sample of the billet's cross-section. A typical bright macrostructure is observable due to the operation of the mould's electromagnetic stirrers (indicated by arrows). Subsurface cracks outside the corner are also visible in the upper left and lower right corners. The same billet macrostructure and open cracks outside the corners can be seen in the cross-section of the macro-etched round rolled bar, as shown in Fig. 2.

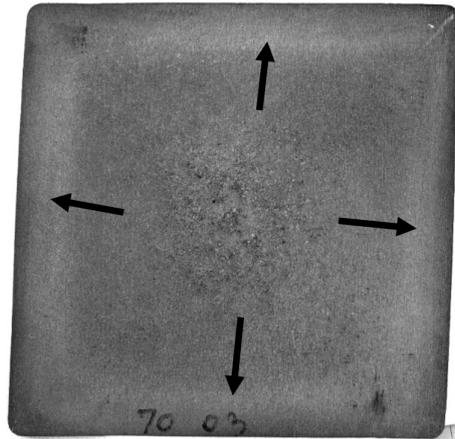


Fig. 1 A macro-etched sample of the billet's cross-section reveals a typical bright macrostructure, which can be observed due to the operation of the mould's electromagnetic stirrers (indicated by arrows). Subsurface cracks outside the corner are also visible in the upper left and lower right corners.

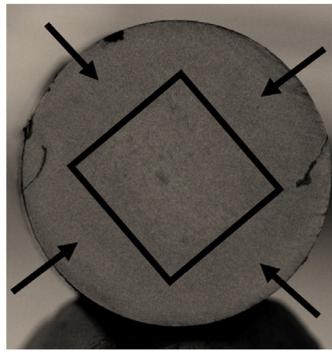


Fig. 2 The black square signifies the orientation of the billet. Both defects (on the left and right) are located at the corner of the billet, indicating the presence of off-corner cracks in the billet, which is a cast semi-product. The operation of the mould's electromagnetic stirrers results in a typically bright macro-structure, as indicated by the arrows.

In May 2023, a project was started where the casting speed was adjusted according to the casting temperature. This adjustment involved steel grades with the highest number of surface defect occurrences (i.e., the average ratio of rolled bars with detected surface defects and their lengths) and varying carbon contents: 16MnCrS5, C22, 30MnVS5, and 46MnVS5. For every 10 °C deviation from the prescribed casting temperature, the speed was altered by 0.02 m/min. The number of cast batches, the average ratio of rolled bars with detected surface defects, and their lengths for the one-year period from May 2022 to May 2023 are presented in Table 1.

In an effort to reduce casting defects on the rolled material, the following parameters were collected for the batches cast from May 2023 to June 2023. This two-month period involved adjustments to the casting speed based on the casting temperature. The adjustments were specifically applied to steel grades with the highest occurrences of surface defects, namely 16MnCrS5, C22, 30MnVS5, and 46MnVS5:

- Chemical composition: Content of carbon, silicon, manganese, sulphur, chromium, nickel, aluminium and vanadium. Chemical composition influence on material properties also during solidification (e.g., shrinkage, ductility, mechanical properties).
- Casting parameters:
 - Average casting temperature (in °C). Casting temperature influences the thermal field in the mould, which influences the heat removal and solidification.
 - Changes of the casting speed based on deviation from the prescribed casting temperature (m/min). For every 10 °C deviation from the prescribed casting temperature, the speed was changed by 0.02 m/min. Casting speed influences the heat removal and solidification.

- The average mould water flow (in l/min). The highest heat removal occurs in the mould, where thermomechanical behaviour influences on shell solidification.
- The average cooling water pressure (in bar) and flow (in l/min) are observed in the first zone of secondary cooling, which is directly below the mould. The melt primarily solidifies in the mould. After exiting the mould, the strand is cooled by water sprays. The water flux can be automatically adjusted, varying the water pressure and flow. Secondary cooling also influences the thermomechanical behaviour during solidification.
- Average ratio of rolled bars with detected surface defects (in %).
- Average length of detected surface defects (in mm/mm).

Rolled bars are examined using an automatic control line equipped with a flux leakage inspection system, which has a surface defect depth detection limit of 0.15 mm. Data on defect depths and lengths for each examined bar are available. It's important to note that scrap is considered when the maximum permissible depth of surface defects is exceeded. The maximum permissible depth is defined by the customer or international standards (e.g., ISO 9443, EN 10221, EN 10277-1) and can be significantly (i.e., several times) larger than the detection limit. Accordingly, this study uses data on detectable defects that are deeper than the detection limit of 0.15 mm.

Table 1 The number of cast batches, average ratio of rolled bars with detected surface defects and their lengths for individual steel grade for the one-year period from May 2022 and May 2023

Steel grade	Number of cast batches	Average ratio of rolled bars with the detected surface defects, %	The length of the detected surface defects, mm/m
C45S	255	12.92	10.57
46MNV5	67	25.52	18.08
16MNCR5	70	11.27	1.41
30MNV6	29	16.28	12.29
42CRM04	26	13.90	3.73
28MNCRNIB	43	6.97	4.88
20MNCR5	43	5.65	0.30
C45	36	6.36	3.07
20MNV6	83	2.62	0.57
C22	17	12.73	1.425
S355J2	28	7.59	1.46
42CRM04	23	8.79	1.55
51CRV4	22	8.52	3.48
20NICRMOS2-2	16	8.26	0.63
16MNCR5	28	3.81	0.47
23MNNICRM05-2	15	7.04	8.80
18CRNIM07-6	18	5.84	0.07
C35S	7	11.38	32.62
25CRM04	6	10.91	9.88
28MNCRB7	6	9.15	1.80
25CRM04	8	6.30	0.06
100CR6	9	5.02	2.13
C60	19	2.25	0.26
20MNCR5	6	6.04	0.27
31CRM0V9	4	6.67	2.71
38MNV6	7	3.34	0.50
16NICRS4	7	3.06	6.62
17NICRMOS6-4	4	2.09	0.08
30CRNIM08	3	2.11	2.67
15CRNI6	3	2.06	0.40
P460NH	3	1.38	0.09

Before implementing changes to the casting speed based on deviations from the prescribed casting temperature into the continuous casting process, models were developed using linear regression and genetic programming. These models predict the ratio of rolled bars with detected surface defects and the length of these defects. By comparing the calculated ratios of rolled bars with detected surface defects and their lengths, with or without changes in casting speed, further measures were taken.

Table 2 Data from May 2023 to June 2023 where the casting temperature has been adjusted, including steel grades with the highest number of surface defect occurrences – 16MnCrS5, C22, 30MnVS5 and 46MnVS5

Batch number	Steel grade	C C, %	Si Si, %	Mn Mn, %	S S, %	Cr Cr, %	Ni Ni, %	Al Al, %	V V, %	Changes of the casting speed SPEED, m/min	The average cooling water flow in the mould FLUXM, l/min	The average cooling water flow in the first zone of secondary cooling FLUXZ1, l/min	The average cooling water pressure in the first zone of secondary cooling PRESS, bar	The average casting temperature TEMP, °C	Average ratio of rolled bars with the detected surface defects, %	The length of the de- tected surface defects, mm/m
1	16MnCrS5	0.16	0.24	1.11	0.009	1.00	0.19	0.021	0.01	0.00	1780.03	35.21	2.42	1553.0	0.58	0.12
2	16MnCrS5	0.16	0.24	1.12	0.008	1.02	0.17	0.022	0.01	0.00	1780.01	35.19	2.42	1552.0	1.80	0.11
3	16MnCrS5	0.16	0.25	1.12	0.008	1.00	0.13	0.022	0.01	0.02	1780.04	34.99	2.55	1564.0	0.59	0.12
4	16MnCrS5	0.17	0.25	1.12	0.007	0.99	0.16	0.022	0.01	0.00	1780.05	34.84	2.43	1551.0	2.27	0.17
5	30MnVS6	0.29	0.58	1.43	0.027	0.12	0.09	0.013	0.09	0.02	1780.01	35.09	2.41	1542.0	2.91	0.67
6	30MnVS6	0.30	0.60	1.40	0.033	0.12	0.07	0.016	0.08	0.00	1780.03	35.01	2.44	1539.0	2.44	1.03
7	30MnVS6	0.29	0.56	1.41	0.033	0.20	0.12	0.013	0.09	0.02	1779.99	35.14	2.49	1542.0	5.56	0.21
8	30MnVS6	0.30	0.60	1.37	0.026	0.19	0.10	0.013	0.10	0.00	1780.03	34.97	2.47	1538.0	0.97	0.09
9	46MnVS5	0.47	0.65	1.15	0.062	0.26	0.15	0.006	0.11	0.00	1780.03	32.26	2.30	1517.0	3.45	1.22
10	46MnVS5	0.47	0.63	1.17	0.063	0.26	0.16	0.006	0.11	0.02	1780.02	32.11	2.28	1524.0	3.52	1.32
11	46MnVS5	0.47	0.64	1.15	0.068	0.24	0.16	0.005	0.11	0.02	1779.97	31.94	2.00	1513.0	2.45	0.31
12	46MnVS5	0.47	0.65	1.15	0.063	0.24	0.16	0.005	0.11	0.02	1780.03	32.53	2.08	1525.0	11.16	1.41
13	46MnVS5	0.47	0.64	1.15	0.063	0.24	0.16	0.005	0.11	0.00	1780.01	32.10	2.03	1514.0	21.10	2.16
14	46MnVS5	0.47	0.66	1.15	0.066	0.25	0.16	0.005	0.11	0.02	1780.01	31.81	1.99	1516.0	18.07	2.36
15	46MnVS5	0.47	0.66	1.15	0.060	0.25	0.17	0.005	0.11	0.02	1780.02	32.25	2.04	1534.0	12.48	4.85
16	46MnVS5	0.47	0.64	1.15	0.067	0.24	0.17	0.005	0.11	0.00	1780.00	31.72	1.98	1526.0	17.58	9.21
17	46MnVS5	0.47	0.64	1.15	0.064	0.24	0.16	0.005	0.11	0.00	1780.00	31.96	2.01	1513.0	16.28	19.68
18	46MnVS5	0.47	0.66	1.14	0.063	0.24	0.16	0.005	0.11	0.00	1780.00	31.90	2.11	1516.0	2.22	2.07
19	C22	0.22	0.21	0.43	0.006	0.20	0.11	0.020	0.00	0.00	1780.03	42.22	3.50	1554.0	2.43	0.14
20	C22	0.22	0.24	0.41	0.007	0.14	0.10	0.019	0.00	0.02	1779.98	42.16	3.49	1558.0	1.45	0.04
21	C22	0.22	0.23	0.44	0.006	0.10	0.07	0.018	0.00	0.02	1779.95	42.15	3.48	1555.0	1.39	0.12
22	C22	0.22	0.25	0.43	0.003	0.08	0.08	0.024	0.00	0.02	1779.99	42.20	3.49	1552.0	1.82	0.64
23	C22	0.23	0.24	0.43	0.004	0.11	0.08	0.020	0.00	0.02	1780.02	42.18	3.48	1553.0	1.55	0.18
24	C22	0.22	0.25	0.46	0.008	0.14	0.09	0.020	0.00	0.00	1780.00	42.17	3.48	1555.0	2.18	0.12

3. Results and discussion

Based on the collected data (Table 2), the prediction of the average ratio of rolled bars with detected surface defects, as well as the average length of detected surface defects, was conducted using linear regression and genetic programming. The fitness function was defined as the average deviation between the predicted and experimental data. It is defined as follows:

$$\Delta = \frac{\sum_{i=1}^n |X_i - X'_i|}{n} \tag{1}$$

where n is the size of the monitored data and X'_i and X_i are the actual and the predicted (i.e. calculated) values, respectively.

3.1 Modelling of the ratio of rolled bars with detected surface defects and their lengths using linear regression

Linear regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables. It provides a way to predict the dependent variable's value based on the values of the independent variables, making it a valuable tool in fields such as machine learning, economics, engineering, and biology.

Based on the linear regression results, it can be concluded that the model significantly predicts the average ratio of rolled bars with detected surface defects ($p < 0.05$, ANOVA). It is found that 72.78 % of total variances can be explained by the variances of independent variables (R-square). The only significantly influential parameter is the average cooling water pressure in the first zone of secondary cooling ($PRESS$) ($p < 0.05$).

The linear regression model for predicting the average ratio of rolled bars with detected surface defects is as follows:

$$\begin{aligned}
 & -0.789 \cdot C - 0.450 \cdot Si - 1.357 \cdot Mn + -0.235 \cdot S + 3.577 \cdot Cr - 0.302 \cdot Ni - 0.576 \cdot Al + 2.822 \\
 & \cdot V + 3.772 \cdot SPEED + 0.711 \cdot FLUXM + 0.015 \cdot FLUXZ1 - 0.455 \cdot PRESS \quad (2) \\
 & + 0.002 \cdot TEMP - 1268.557.
 \end{aligned}$$

The average deviation from experimental data is 2.32 %.

Similarly, the second model significantly predicts the average length of detected surface defects ($p < 0.05$, ANOVA). However, in this case, only 50.17 % of total variances can be explained by the variances of independent variables (R-square). Interestingly, in this context, there are no significantly influential parameters ($p < 0.05$).

The linear regression model for predicting the average length of detected surface defects is as follows:

$$\begin{aligned}
 & -179.589 \cdot C + 44.737 \cdot Si - 116.991 \cdot Mn - 7.780 \cdot S - 99.020 \cdot Cr - 32.913 \cdot Ni - 39.153 \cdot Al + \\
 & 409.933 \cdot V + 188.939 \cdot SPEED - 22.826 \cdot FLUXM - 2.781 \cdot FLUXZ1 - 17.145 \cdot PRESS + 0.228 \cdot \quad (3) \\
 & TEMP + 40476.188.
 \end{aligned}$$

The average deviation from the experimental data is 5.44 mm/m.

Based on the data gathered in Table 2 and the developed linear regression models, the average ratio of rolled bars with detected surface defects and the average length of detected surface defects were also calculated in the scenario where changes to the casting speed were not made.

According to calculations from both linear regression models, changes in casting speed contributed to a decrease in the average ratio of rolled bars with detected surface defects and the average length of detected surface defects, from 6.47 % to 5.68 % and from 3.81 mm/m to 2.01 mm/m, respectively.

3.2 Modelling of the ratio of rolled bars with detected surface defects and their lengths using genetic programming

Genetic programming is an evolutionary algorithm, similar to genetic algorithms, used for automatic generation of computer programs to solve problems. It involves evolving a population of computer programs over several generations, using genetic operators like crossover and mutation to produce new candidate solutions. The programs are represented as tree structures, allowing for the evolution of complex solutions. In genetic programming, the representation of solutions as tree structures enables the evolution of diverse and complex programs, allowing for the exploration of a broad solution space. Unlike genetic algorithms which typically evolve fixed-length strings, which are intended to solve a very broad spectrum of problems [22-28], genetic programming evolves variable-sized structures, allowing for more flexibility in representing solutions of varying complexity [29-31]. Genetic programming can automatically discover both the structure and parameters of a solution, making it suitable for problems where the optimal solution's form is not known a priori. The genetic programming operates on variable-length structures, making it more suitable for evolving complex solutions, especially in symbolic regression and automatic code generation.

The genetic programming method was used several times in Štore Steel Ltd. [32-36]. For the purposes of this study, organisms that underwent adaptation were indeed represented as mathematical expressions, i.e., models for predicting the average ratio of rolled bars with detected surface defects, and models for predicting the average length of detected surface defects. These models consist of the selected functions, i.e., basic arithmetical functions of addition, subtraction, multiplication and division, and terminal genes, i.e., independent input parameters, and random floating-point constants.

The LISP based in-house genetic programming system was run 200 times to develop independent civilizations. In each run we obtain either the model for prediction of the average ratio of rolled bars with detected surface defects or the average length of detected surface defects. After the modelling phase, we analysed the results and selected the two best prediction models.

Table 3 Data from May 2023 to June 2023 where the casting temperature has been adjusted, including steel grades with the highest number of surface defect occurrences - 16MnCrS5, C22, 30MnVS5 and 46MnVS5

Steel grade	Number of batches from May 2022 to May 2023 (period without casting speed adjustments)	Number of batches from July to November 2023 (period with casting speed adjustments)	Number of batches from July to November 2023 (period with casting speed adjustments), where the casting speed was actually adjusted	The average ratio of rolled bars with detected surface defects from May 2022 to May 2023 (period without casting speed adjustments), %	The average ratio of rolled bars with detected surface defects from July to November 2023 (period with casting speed adjustments), %	The average length of detected surface defects from May 2022 to May 2023 (period without casting speed adjustments), mm/m	The average length of detected surface defects from July to November 2023 (period with casting speed adjustments), mm/m	The average ratio of rolled bars with detected surface defects – significance (t-test, $p < 0.05$)	The average length of detected surface defects – significance (t-test, $p < 0.05$)
17NICRMOS6-4	4	3	3	2.09	4.62	0.18	0.08	NO	NO
30CRNIM08	3	2	2	2.11	0.56	2.67	0.02	NO	NO
C60	19	9	6	2.25	3.47	0.26	0.83	NO	NO
20MNV6	83	39	23	2.62	1.29	0.57	0.19	YES	YES
16NICRS4	7	4	3	3.06	2.79	6.62	0.21	NO	YES
38MNV56	7	6	2	3.34	6.88	0.50	1.03	NO	NO
16MNCRS5	28	12	5	3.81	4.01	1.31	0.47	NO	YES
100CR6	9	2	0	5.02	10.43	2.13	1.98	NO	NO
18CRNIM07-6	18	7	4	5.84	5.89	0.07	0.05	NO	NO
20MNCRS5	6	2	1	6.04	3.19	0.27	0.35	NO	NO
25CRM04	8	8	5	6.30	4.84	0.06	0.09	NO	NO
C45	36	18	16	6.36	1.95	3.07	0.32	YES	YES
31CRMOV9	4	4	3	6.67	7.46	2.71	2.82	NO	NO
28MNCRNIB	43	8	4	6.97	5.40	4.87	1.27	YES	YES
15CRNI6	3	2	1	6.98	2.06	1.48	0.40	YES	YES
23MNNICRM05-2	15	9	3	7.04	6.50	8.80	0.51	NO	YES
S35J2	28	10	4	7.59	3.78	1.46	0.76	YES	NO
51CRV4	22	4	4	8.52	12.18	3.48	0.62	NO	YES
42CRM04	23	5	2	8.79	10.54	1.55	0.96	NO	NO
28MNCRB7	6	4	2	9.15	16.04	1.80	1.55	NO	NO
P460NH	3	4	2	9.23	1.38	0.75	0.09	YES	YES
20MNCRS5	43	15	1	10.59	5.65	1.12	0.30	YES	NO
25CRMOS4	6	3	2	10.91	6.00	9.88	0.57	NO	YES
16MNCRS5	70	39	9	11.27	7.93	1.41	1.29	YES	NO
C35S	7	4	2	11.38	5.61	32.62	6.00	YES	NO
C22	17	18	9	12.73	4.11	1.43	0.15	YES	YES
C45S	255	141	65	12.92	13.00	10.57	6.46	NO	YES
20NICRMOS2-2	16	9	3	13.45	8.26	1.83	0.63	NO	NO
42CRMOS4	26	10	5	13.90	9.71	3.73	3.28	NO	NO
30MNV56	29	12	5	16.28	13.40	12.29	4.57	NO	NO
46MNV55	67	35	4	25.52	16.99	18.08	9.98	YES	YES

4. Conclusion

In the article the reduction of surface defects, detected during examination (automatic control line) of the rolled material, with casting speed adjustments based on the casting temperature, is presented. For all produced steel grades, the same casting speed of 1.6 m/min was used before. The average casting speed adjustments were calculated for several different grades based on in-house developed solidification model. For every 10 °C deviation from the prescribed casting temperature, the speed was changed by 0.02 m/min.

The reduction of surface defects was designed as follows:

- Period with casting speed adjustments based on the casting temperature involving steel grades with highest occurrences of surface defects and with various carbon content: 16MnCrS5, C22, 30MnVS5 and 46MnVS5.
- Modelling of the average ratio of rolled bars with detected surface defects and the average length of detected surface defects using linear regression and genetic programming.
- Implementing of modelling results based on four most problematic steel grades (16MnCrS5, C22, 30MnVS5 and 46MnVS5) into practice – the casting speed adjustments were used for all steel grades from July 2023 to November 2023.

- Analysis of the average ratio of rolled bars with detected surface defects and the average length of detected surface defects for batches cast from July 2023 to November 2023.
- Comparison of the batches produced prior (from May 2022 to May 2023) and after (from July 2023 to November 2023) the implementation of casting speed adjustments.

Chemical composition (content of carbon, silicon, manganese, sulphur, chromium, nickel, aluminium and vanadium) and casting parameters (casting speed adjustments, the average mould cooling water flow, the average cooling water flow and pressure in the first zone of the secondary cooling, casting temperature) were gathered for the batches cast from May 2023 to June 2023.

Based on gathered data modelling was performed using linear regression and genetic programming. For the fitness function, the average deviation between predicted and experimental data was selected.

The average deviation of the linear regression model for predicting the average ratio of bars with detected surface defects from experimental data is 2.32 %. The average deviation of the linear regression model for predicting the average length of detected surface defects from experimental data is 5.44 mm/m. Based on calculations from both linear regression models, the changes in casting speed contributed to a decrease in the average ratio of rolled bars with detected surface defects and the average length of detected surface defects, from 6.47 % to 5.68 % and 3.81 mm/m to 2.01 mm/m, respectively.

The average deviation of the genetic programming model for predicting the average ratio of bars with detected surface defects from experimental data is 1.59 %. The model obtained by genetic programming is 1.47 times better than the one obtained using linear regression. The average deviation of the linear regression model for predicting the average length of detected surface defects from experimental data is 0.97 mm/m. The model obtained by genetic programming is 5.63 times better than the one obtained using linear regression. Based on calculations using both obtained genetic programming models, the changes of casting speed contributed to decreasing the average ratio of rolled bars with detected surface defects and the average length of detected surface defects, from 6.1 % to 6.0 % and 8.87 mm/m to 1.57 mm/m, respectively.

Accordingly, the casting speed has been adjusted based on the casting temperature for all cast batches from July 2023 to November 2023. The obtained data were compared with a one-year period (from May 2022 to May 2023) where adjustments to the casting speed were not made. Based on actual data, the average ratio of rolled bars with detected surface defects and the average length of detected surface defects decreased statistically significantly (t-test, $p < 0.05$) from 10.63 % to 8.82 % and from 6.20 mm/m to 3.40 mm/m in the period without (from May 2022 to May 2023) and with casting speed adjustments (from July 2023 to November 2023), respectively. Out of 31 different steel grades, the average ratio of rolled bars with detected surface defects and the average length of detected surface defects statistically significantly decreased in 11 and 15 steel grades, respectively, while the rest remained statistically insignificantly the same (t-test, $p < 0.05$).

Until June 2024, detailed analyses of the possible casting speed adjustments for individual steel grades will be conducted with the in-house developed solidification model. The geometry of the mould and the geometry, pumps, and nozzles of the secondary cooling will be changed.

References

- [1] Emi, T., Fredriksson, H. (2005). High-speed continuous casting of peritectic carbon steels, *Materials Science and Engineering: A*, Vol. 413-414, 2-9, doi: [10.1016/j.msea.2005.08.169](https://doi.org/10.1016/j.msea.2005.08.169).
- [2] Santos, C.A., Spim, J.A., Ierardi, M.C., Garcia, A. (2002). The use of artificial intelligence technique for the optimisation of process parameters used in the continuous casting of steel, *Applied Mathematical Modelling*, Vol. 26, No. 11, 1077-1092, doi: [10.1016/S0307-904X\(02\)00062-8](https://doi.org/10.1016/S0307-904X(02)00062-8).
- [3] Boonpen, K., Kowitwarangkul, P., Ninpetch, P., Phophichit, N., Chuchuy, P., Threrujirapapong, T., Otarawanna, S. (2021). Numerical study of influence of casting speed on fluid flow characteristics in the four strand tundish, *Materials Today: Proceedings*, Vol. 47, Part 12, 3480-3486, doi: [10.1016/j.matpr.2021.03.465](https://doi.org/10.1016/j.matpr.2021.03.465).

- [4] Hore, S., Das, S.K., Humane, M.M., Peethala, A.K. (2019). Neural network modelling to characterize steel continuous casting process parameters and prediction of casting defects, *Transactions of the Indian Institute of Metals*, Vol. 72, No. 12, 3015-3025, doi: [10.1007/s12666-019-01767-0](https://doi.org/10.1007/s12666-019-01767-0).
- [5] Fei, P., Min, Y., Liu, C.-J., Jiang, M.-F. (2019). Effect of continuous casting speed on mold surface flow and the related near-surface distribution of non-metallic inclusions, *International Journal of Minerals, Metallurgy, and Materials*, Vol. 26, No. 2, 186-193, doi: [10.1007/s12613-019-1723-y](https://doi.org/10.1007/s12613-019-1723-y).
- [6] Swain, A.N.S.S., Ganguly, S., Sengupta, A., Chacko, E.Z., Dhakate, S., Pandey, P.K. (2022). Investigation of corner cracks in continuous casting billet using thermomechanical model and plant measurements, *Metals and Materials International*, Vol. 28, No. 10, 2434-2447, doi: [10.1007/s12540-021-01135-y](https://doi.org/10.1007/s12540-021-01135-y).
- [7] Zhu, L.-G., Kumar, R.V. (2007). Modelling of steel shrinkage and optimisation of mould taper for high speed continuous casting, *Ironmaking & Steelmaking*, Vol. 34, No. 1, 76-82, doi: [10.1179/174328106X118152](https://doi.org/10.1179/174328106X118152).
- [8] Alizadeh, M. (2015). Study on hot tearing tendency during continuous casting of steel by overall hot tearing susceptibility (OHTS), *International Journal of Cast Metals Research*, Vol. 28, No. 1, 20-27, doi: [10.1179/1743133614Y.0000000124](https://doi.org/10.1179/1743133614Y.0000000124).
- [9] Furumai, K., Aramaki, N., Oikawa, K. (2022). Influence of heat flux different between wide and narrow face in continuous casting mould on unevenness of hypo-peritectic steel solidification at off-corner, *Ironmaking & Steelmaking*, Vol. 49, No. 9, 845-859, doi: [10.1080/03019233.2022.2063654](https://doi.org/10.1080/03019233.2022.2063654).
- [10] Sala, D.A., Van Yperen-De Deyne, A., Mannens, E., Jalalvand, A. (2023). Hybrid-input FCN-CNN-SE for industrial applications: Classification of longitudinal cracks during continuous casting, *Metals*, Vol. 13, No. 10, Article No. 1699, doi: [10.3390/met13101699](https://doi.org/10.3390/met13101699).
- [11] Kong, Y., Chen, D., Liu, Q., Long, M. (2019). A prediction model for internal cracks during slab continuous casting, *Metals*, Vol. 9, No. 5, Article No. 587, doi: [10.3390/met9050587](https://doi.org/10.3390/met9050587).
- [12] Zhang, T., Yang, J., Jiang, P. (2019). Measurement of molten steel velocity near the surface and modeling for transient fluid flow in the continuous casting mold, *Metals*, Vol. 9, No. 1, Article No. 36, doi: [10.3390/met9010036](https://doi.org/10.3390/met9010036).
- [13] Kovačič, M., Jager, R. (2015). Modeling of occurrence of surface defects of C45 steel with genetic programming, *Materiali in tehnologije/Materials and technology*, Vol. 49, No. 6, 857-863, doi: [10.17222/mit.2013.304](https://doi.org/10.17222/mit.2013.304).
- [14] Tran, H.-S., Castiaux, E., Habraken, A.-M. (2020). 2D thermal finite element analysis of sticker breakout in continuous casting, *Procedia Manufacturing*, Vol. 50, 376-383, doi: [10.1016/j.promfg.2020.08.069](https://doi.org/10.1016/j.promfg.2020.08.069).
- [15] He, F., Zhou, L., Deng, Z.-H. (2015). Novel mold breakout prediction and control technology in slab continuous casting, *Journal of Process Control*, Vol. 29, 1-10, doi: [10.1016/j.jprocont.2015.03.003](https://doi.org/10.1016/j.jprocont.2015.03.003).
- [16] Santos, C.A., Spim, J.A., Ierardi, M.C.F., Garcia, A. (2002). The use of artificial intelligence technique for the optimization of process parameters used in the continuous casting of steel, *Applied Mathematical Modelling*, Vol. 26, No. 11, 1077-1092, doi: [10.1016/S0307-904X\(02\)00062-8](https://doi.org/10.1016/S0307-904X(02)00062-8).
- [17] Tao, J., Wang, N. (2005). Fuzzy neuron hybrid control for continuous steel casting, *IFAC Proceedings Volumes*, Vol. 38, No. 1, 121-126, doi: [10.3182/20050703-6-CZ-1902.01699](https://doi.org/10.3182/20050703-6-CZ-1902.01699).
- [18] Zhang, Q.-Y., Wang, X.-H. (2010). Numerical simulation of influence of casting speed variation on surface fluctuation of molten steel in mold, *Journal of Iron and Steel Research International*, Vol. 17, No. 8, 15-19, doi: [10.1016/S1006-706X\(10\)60121-5](https://doi.org/10.1016/S1006-706X(10)60121-5).
- [19] Vertnik, R., Mramor, K., Šarler, B. (2019). Solution of three-dimensional temperature and turbulent velocity field in continuously cast steel billets with electromagnetic stirring by a meshless method, *Engineering Analysis with Boundary Elements*, Vol. 104, 347-363, doi: [10.1016/j.enganabound.2019.03.026](https://doi.org/10.1016/j.enganabound.2019.03.026).
- [20] Šarler, B., Vertnik, R., Šaletić, S., Manojlović, G., Cesar, J. (2005). Application of continuous casting simulation at štore steel, *BHM Berg- und Hüttenmännische Monatshefte*, Vol. 150, No. 9, 300-306, doi: [10.1007/BF03165327](https://doi.org/10.1007/BF03165327).
- [21] Gryc, K., Smetana, B., Žaludová, M., Michalek, K., Klus, P., Tkadlečková, M., Socha, L., Dobrovská, J., Machovčák, P., Válek, L., Pachlopnik, R., Chmiel, B. (2013). Determination of the solidus and liquidus temperatures of the real-steel grades with dynamic thermal-analysis methods, *Materiali in tehnologije/Materials and technology*, Vol. 47, No. 5, 565-575.
- [22] Bhoskar, T., Kulkarni, O.K., Kulkarni, N.K., Patekar, S.L., Kakandikar, G.M., Nandedkar, V.M. (2015). Genetic algorithm and its applications to mechanical engineering: A review, *Materials Today: Proceedings*, Vol. 2, No. 4-5, 2624-2630, doi: [10.1016/j.matpr.2015.07.219](https://doi.org/10.1016/j.matpr.2015.07.219).
- [23] Jha, R., Pettersson, F., Dulikravich, G.S., Saxen, H., Chakraborti, N. (2015) Evolutionary design of nickel-based superalloys using data-driven genetic algorithms and related strategies, *Materials and Manufacturing Processes*, Vol. 30, No. 4, 488-510, doi: [10.1080/10426914.2014.984203](https://doi.org/10.1080/10426914.2014.984203).
- [24] Gajsek, B., Dukic, G., Kovacic, M., Brezocnik, M. (2021). A multi-objective genetic algorithms approach for modeling of order picking, *International Journal of Simulation Modelling*, Vol. 20, No. 4, 719-729, doi: [10.2507/IJSIMM20-4-582](https://doi.org/10.2507/IJSIMM20-4-582).
- [25] Tuzkaya, U.R., Şahin, S. (2021). A single side priority based ga approach for 3D printing center integration to spare part supply chain in automotive industry, *Tehnički Vjesnik – Technical Gazette*, Vol. 28, No. 3, 836-844, doi: [10.17559/TV-20200311104539](https://doi.org/10.17559/TV-20200311104539).
- [26] Liu, M.L., Yao, X.Z., Huang, J.Y., Zhang, C. (2022). Optimization of unmanned vehicle scheduling and order allocation, *International Journal of Simulation Modelling*, Vol. 21, No. 3, 477-488, doi: [10.2507/IJSIMM21-3-613](https://doi.org/10.2507/IJSIMM21-3-613).
- [27] Chen, D., Zhao, X.R. (2021). Production management of hybrid flow shop based on genetic algorithm, *International Journal of Simulation Modelling*, Vol. 20, No. 3, 571-582, doi: [10.2507/IJSIMM20-3-C012](https://doi.org/10.2507/IJSIMM20-3-C012).
- [28] Grubeša, S., Suhaneck, M., Djurek, I., Petošić, A. (2021). Combination of boundary element method and genetic algorithm for optimization of T-shape noise Barrier, *Tehnički Vjesnik – Technical Gazette*, Vol. 28, No. 1, 77-81, doi: [10.17559/TV-20190930132137](https://doi.org/10.17559/TV-20190930132137).

- [29] Koza, J.R. (1992). *Genetic programming: On the programming of computers by means of natural selection*, MIT Press Cambridge, USA.
- [30] Koza, J.R. (1994). *Genetic programming II: Automatic discovery of reusable programs*, MIT Press, Cambridge, USA.
- [31] Koza, J.R., Andre, D., Bennett, F.H., Keane, M.A. (1999). Genetic programming III: Darwinian invention and problem solving [book review], *IEEE Transactions on Evolutionary Computation*, Vol. 3, No. 3, 251-253, [doi: 10.1109/TEVC.1999.788530](https://doi.org/10.1109/TEVC.1999.788530).
- [32] Kovacic, M., Brezocnik, M. (2018). Reduction of surface defects and optimization of continuous casting of 70MnVS4 steel, *International Journal of Simulation Modelling*, Vol. 17, No. 4, 667-676, [doi: 10.2507/IJSIMM17\(4\)457](https://doi.org/10.2507/IJSIMM17(4)457).
- [33] Kovačič, M., Župerl, U. (2023). Continuous caster final electromagnetic stirrers position optimization using genetic programming, *Materials and Manufacturing Processes*, Vol. 38, No. 16, 2009-2017, [doi: 10.1080/10426914.2023.2219317](https://doi.org/10.1080/10426914.2023.2219317).
- [34] Kovačič, M., Župerl, U. (2022). Modeling of tensile test results for low alloy steels by linear regression and genetic programming taking into account the non-metallic inclusions, *Metals*, Vol. 12, No. 8, Article No. 1343, [doi: 10.3390/met12081343](https://doi.org/10.3390/met12081343).
- [35] Kovačič, M., Šarler, B. (2011). Genetic programming and soft-annealing productivity, *Materiali in tehnologije/Materials and technology*, Vol. 45, No. 5, 427-432.
- [36] Kovačič, M., Župerl, U. (2020). Genetic programming in the steelmaking industry, *Genetic Programming and Evolvable Machines*, Vol. 21, No. 1-2, 99-128, [doi: 10.1007/s10710-020-09382-5](https://doi.org/10.1007/s10710-020-09382-5).

Calendar of events

- European Simulation and Modelling Conference (ESM 2023), October 24-26, 2023, Toulouse, France.
- 17th International Conference on Industrial and Manufacturing Systems Engineering, November 27-28, 2023, London, United Kingdom.
- International Conference on Digital Manufacturing and Industrial Design ICDMID on December 13-14, 2023, Rome, Italy.
- International Conference on Digital Manufacturing and Automation ICDMA 2023, December 13-15, 2023, Singapore, Singapore.
- International Conference on Multiple Machines and Manufacturing Systems ICMMMS, December 18-19, 2023, Bangkok, Thailand.
- International Conference on Design and Development of Reconfigurable Manufacturing Systems ICDDRMS, December 20-21, 2023, Honolulu, USA.
- International Conference on Digital Manufacturing, Modeling and Design ICDMMD, December 20-21, 2023, Dubai, United Arab Emirates.
- International Conference on Digital Factory and Manufacturing ICDFM, December 20-21, 2023, Dubai, United Arab Emirates.
- International Conference on Intelligent Manufacturing and Automation Engineering ICIMA, December 22-24, 2023, Changsha, China.
- 17. International Conference on Rapid Manufacturing December 25-26, 2023, Paris, France.
- International Conference on Digital Manufacturing Systems ICDMS, January 7-8, 2024, Tokyo, Japan.
- International Conference on Manufacturing and Intelligent Machining ICMIM, January 14-15, 2024, Zurich, Switzerland.
- 18th International Conference on Advanced Manufacturing Engineering and Technologies, January 15-16, 2024, Montevideo, Uruguay.
- International Conference on Cyber Manufacturing Systems ICCMS, January 18-19, 2024, Sydney, Australia.
- International Conference on Manufacturing and Industrial Technologies ICMIT, January 25-27, 2024, Budapest, Hungary.
- International Conference on Mechatronics and Manufacturing ICMM, February 2-4, 2024, Penang, Malaysia.
- International Conference on Manufacturing and Optimization ICMO, February 25-26, 2024, Buenos Aires, Argentina.
- International Conference on Mechanical and Intelligent Manufacturing Technologies ICMIMT, March 7-9, 2024, Cape Town, South Africa.
- 2024 Annual Modeling and Simulation Conference (ANNSIM 2024), May 20-23, 2024, Washington D.C., USA.
- North American manufacturing research conference (NAMRC) 52, June 17-21, 2024, Knoxville, TN, USA.
- 18th International Conference on Industrial and Manufacturing Systems Engineering, August 9-10, 2024, Lagos, Nigeria.

Notes for contributors

General

Articles submitted to the *APEM journal* should be original and unpublished contributions and should not be under consideration for any other publication at the same time. Manuscript should be written in English. Responsibility for the contents of the paper rests upon the authors and not upon the editors or the publisher. The content from published paper in the *APEM journal* may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). For most up-to-date information please see the APEM journal homepage apem-journal.org.

Submission of papers

A submission must include the corresponding author's complete name, affiliation, address, phone and fax numbers, and e-mail address. All papers for consideration by *Advances in Production Engineering & Management* should be submitted by e-mail to the journal Editor-in-Chief:

Miran Brezocnik, Editor-in-Chief
UNIVERSITY OF MARIBOR
Faculty of Mechanical Engineering
Chair of Production Engineering
Smetanova ulica 17, SI – 2000 Maribor
Slovenia, European Union
E-mail: editor@apem-journal.org

Manuscript preparation

Manuscript should be prepared in *Microsoft Word 2010* (or higher version) word processor. *Word .docx* format is required. Papers on A4 format, single-spaced, typed in one column, using body text font size of 11 pt, should not exceed 12 pages, including abstract, keywords, body text, figures, tables, acknowledgements (if any), references, and appendices (if any). The title of the paper, authors' names, affiliations and headings of the body text should be in *Calibri* font. Body text, figures and tables captions have to be written in *Cambria* font. Mathematical equations and expressions must be set in *Microsoft Word Equation Editor* and written in *Cambria Math* font. For detail instructions on manuscript preparation please see instruction for authors in the *APEM journal* homepage apem-journal.org.

The review process

Every manuscript submitted for possible publication in the *APEM journal* is first briefly reviewed by the editor for general suitability for the journal. Notification of successful submission is sent. After initial screening, and checking by a special plagiarism detection tool, the manuscript is passed on to at least two referees. A double-blind peer review process ensures the content's validity and relevance. Optionally, authors are invited to suggest up to three well-respected experts in the field discussed in the article who might act as reviewers. The review process can take up to eight weeks on average. Based on the comments of the referees, the editor will take a decision about the paper. The following decisions can be made: accepting the paper, reconsidering the paper after changes, or rejecting the paper. Accepted papers may not be offered elsewhere for publication. The editor may, in some circumstances, vary this process at his discretion.

Proofs

Proofs will be sent to the corresponding author and should be returned within 3 days of receipt. Corrections should be restricted to typesetting errors and minor changes.

Offprints

An e-offprint, i.e., a PDF version of the published article, will be sent by e-mail to the corresponding author. Additionally, one complete copy of the journal will be sent free of charge to the corresponding author of the published article.

APEM

journal

Advances in Production Engineering & Management

Chair of Production Engineering (CPE)
University of Maribor
APEM homepage: apem-journal.org

Volume 18 | Number 4 | December 2023 | pp 399-514

Contents

Scope and topics	402
A combined genetic algorithm and A* search algorithm for the electric vehicle routing problem with time windows Wang, D.L.; Ding, A.; Chen, G.L.; Zhang, L.	403
Optimizing rock breaking performance: The influence of chamfer on polycrystalline diamond compact (PDC) cutters Ju, P.	417
Dynamic price competition market for retailers in the context of consumer learning behavior and supplier competition: Machine learning-enhanced agent-based modeling and simulation Deng, G.F.	434
IoT-based Deep Learning Neural Network (DLNN) algorithm for voltage stability control and monitoring of solar power generation Shweta, R.; Sivagnanam, S.; Kumar, K.A.	447
Comparing Fault Tree Analysis methods combined with Generalized Grey Relation Analysis: A new approach and case study in the automotive industry Shi, J.L.; Lu, Z.C.; Xu, H.H.; Ren, M.M.; Shu, F.L.	462
Optimizing smart manufacturing systems using digital twin Ojstersek, R.; Javernik, A.; Buchmeister, B.	475
Incentive modeling analysis in engineering applications and projects with stochastic duration time Zhao, J.; Su, J.F.	486
Reduction of surface defects by optimization of casting speed using genetic programming: An industrial case study Kovacic, M.; Zuperl, U.; Gusel, L.; Brezocnik, M.	501
Calendar of events	512
Notes for contributors	513

Published by CPE, University of Maribor



apem-journal.org