ISSN 1854-6250

APEM *journal*

Advances in Production Engineering & Management

Volume 18 | Number 3 | September 2023

University of Maribor

Published by CPE
apem-journal.org

Advances in Production Engineering & Management

Identification Statement

APEM journal

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

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

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

University of Maribor

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

APEM Editorial

OPEN ACCESS

Subsidizer: The journal is subsidized by Slovenian Research Agency

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

University of Maribor
Chair of Production Engineering (CPE)

Advances in Production Engineering & Management

Volume 18 | Number 3 | September 2023 | pp 267–398

Contents

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

Advances in Production Engineering & Management
Volume 18 | Number 3 | September 2023 | pp 271–287
https://doi.org/10.14743/apem2023.3.472
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

An improved multi-objective Wild Horse optimization for the dual-resource-constrained flexible job shop scheduling problem: A comparative analysis with NSGA-II and a real case study

Peng, F.[a,b], **Zheng, L.**[a,*]

[a]Department of Industrial Engineering, Tsinghua University, P.R. China
[b]CRRC Academy Co. Ltd, Beijing, P.R. China

ABSTRACT

The equipment manufacturing industry needs skilled workers to operate a specific set of machines following process specifications. Optimizing machine and worker assignments to achieve maximum efficiency is a critical problem for workshop managers. This paper investigates a multi-objective dual-resource-constrained flexible job shop scheduling problem. An improved wild horse optimization (IWHO) algorithm is developed to simultaneously optimize three objectives: makespan, maximum machine workload, and total machine workload. To evaluate the quality of individuals in multi-objective optimization, the Pareto fast non-dominated sorting method is used, and the crowding distance is calculated. To update the algorithm's solution, the crossover and mutation operations are used. Further, a local neighborhood search strategy is employed to enhance searchability and avoid trapping into the local optima. The benchmark of the flexible job shop scheduling problem is extended to create test instances, and the performance of the suggested IWHO algorithm is evaluated compared with the NSGA-II. The computational results show that the IWHO algorithm provides a non-dominated efficient set within a reasonable running time. Furthermore, a buffers and chain coupler assembly process is designed to analyze the practical value of the IWHO algorithm. The proposed solutions can be used to generate daily schedules for managing machines, workers, and production cycles.

ARTICLE INFO

Keywords:
Dual resource constraints;
Flexible job shop scheduling;
Wild horse optimization;
Local search;
Multi-objective optimization;
NSGA-II;
Benchmark analysis

**Corresponding author:*
lzheng@mail.tsinghua.edu.cn
(Zheng, L.)

Article history:
Received 12 July 2023
Revised 25 October 2023
Accepted 29 October 2023

1. Introduction

Scheduling is crucial for increasing the efficiency of manufacturing resources in contemporary production systems. In the last few decades, numerous scheduling problems, including job shop scheduling problem (JSP) [1], flexible job shop scheduling problem (FJSP) [2-4], etc., have been exhaustively studied against the growing demand for flexible and customized manufacturing, typically characterized by short life cycles, small lot sizes, and changing product mixes [5]. In standard FJSP, only the machine flexibility is taken into account, which means that the operations of the jobs can be performed by more than one machine from the group of available machines [6]. However, physical manufacturing resources are not just machines and also cover human resources, materials, transferring tools, etc. [7]. Multiple resources are introduced into the production systems to make the scheduling scenario more realistic. Dual resources-constrained (DRC) systems, in which the capacity of both machines and human operators is limited, are studied in depth [8]. DRC systems are more complex than systems with a single resource, and the scheduling process must account for numerous additional technical challenges.

In this paper, we investigated a dual-resource constraint flexible job shop scheduling problem (DRCFJSP) motivated by the example of an engineering equipment manufacturer's assembled workshop. With the development of Industry 4.0 technologies, more and more factories have started to promote intelligent manufacturing [9]. In the actual workshop manufacturing system, constraints on production resource factors such as fixtures, transport equipment, and laborers exist in addition to equipment constraints. Due to the specific requirements of the production structure, the engineering equipment assembly workshop requires a substantial amount of labor. Robots are incapable of performing a thorough job of installing electric components into spatially complex curved surfaces. In addition, multiple workers are organized to execute the tasks sequentially. In such a case, the DRCFJSP must be considered when determining the operational worker and assistant machine orders.

In the production process with human-computer interaction, the generated schedule is challenging to execute if the worker factor is not considered or the worker efficiency is regarded as a fixed value [10]. The human factor largely determines the performance of the production system with manual operation, which makes the scheduling decision of the system more challenging due to the variability of personnel [11]. The processing performance of the production system is closely related to the operational efficiency of workers [12]. The production process in an assembled workshop environment is becoming increasingly reliant on the efficient management of skilled workers, and the unreasonable allocation of personnel will also result in rising wage costs. In addition, due to the recurring COVID-19 epidemic, the manufacturing industry has difficulty recruiting more skilled workers. Assigning a limited number of skilled workers to appropriate positions can increase production efficiency and ensure the plan's smooth progression.

The DRCFJSP has been extensively studied using various approaches such as mixed-integer programming, constrained programming, heuristics, and meta-heuristics. However, there is a need to devise practical alternate solutions for this problem. To address this issue, this paper proposes an enhanced version of the Wild Horse Optimization (WHO) algorithm tailored to solve the DRCFJSP. The primary contributions of this research are listed as follows.

- A multi-objective approach is proposed for the DRCFJSP that simultaneously minimizes the makespan, the maximum machine workload, and the total machine workload.
- The wild horse optimization algorithm is discretized using the crossover and mutation operators, and a local search strategy is added to avoid local optima.
- A real-world test instance is provided for assembling the buffers and chain couplers. And the performance of the designed algorithm is also analyzed by solving the practical instance.

2. Literature review

Our research aims to design an alternative method to solve the DRCFJSP. We discussed the closest two sides in the existing works: (i) DRCFJSP; and (ii) the involved wild horse optimization algorithm. For the flexible job shop scheduling problem, the readers can refer to the comprehensive surveys [2, 13].

It is essential to find a solution to the multi-objective flexible job shop scheduling problem (MOFJSP), considering various production resources for supporting the development of the manufacturing industry in intelligence, flexibility, and personalized customization. The MOFJSP problem is a complex NP-hard problem, which is widely used in various fields such as aviation equipment, the shipbuilding industry [14], agricultural machinery manufacturing [15], semiconductor manufacturing [16], etc. Among that, the DRCFJSP is studied widely from multiple perspectives. The DRCFJSP is introduced considering machine capacity and labor capacity and can be regarded as dual resource-constrained (DRC) systems that commonly exist in real-world situations [8]. In the DRCFJSP researches, the resources assignment problem and the operation sequence problem are studied concurrently. Nonetheless, the DRCFJSP is NP-hard in the strong sense since its simplified form, FJSP, is NP-hard.

In that case, various heuristics were designed to resolve the DRCFJSP. Lei and Guo [17] studied DRCFJSP to optimize the makespan based on variable neighborhood search. Then, they investigated a DRC job shop scheduling problem with interval processing time and heterogeneous resources and developed a dynamical neighborhood search algorithm [18]. Besides, simulated annealing (SA) and vibration damping optimization were designed to solve the DRCFJSP [19]. Moreover, some new population-based optimization algorithms were also presented to solve it, including genetic algorithm (GA) [20], fruit fly optimization algorithm [21], bat algorithm [22], etc.

Due to the increase in computer power, constraint programming techniques were developed to solve the DRCFJSP [23]. However, operations can not be processed if the workers are not available or lack the requisite skills. Wu *et al*. [24] considered the worker's learning ability in the DRCFJSP and suggested a genetic algorithm to obtain the optimal solutions. Besides, the loading and unloading time of the fixtures was introduced into the DRCFJSP, and a non-dominated sorting genetic algorithm II was proposed to minimize the makespan and the setup time [25]. Workers ' fatigue cannot be ignored when considering the DRCFJSP in the casting workshop. And Tan *et al*. discussed a fatigue-conscious DRCFJSP with the support of the NSGA-II, minimizing the maximum worker fatigue and makespan [26]. Due to the requirement for product customization and on-time delivery in make-to-order companies, the due date-related criteria (the mean tardiness) were introduced into the DRCFJSP, except for the makespan [27]. Recently, sequencing flexibility has been studied in DRCFJSP for minimizing the triple objectives, including makespan, maximal worker workload, and weighted tardiness [28]. Based on the above analysis, most of the current works focus on developing the heuristics in solving the DRCFJSP. Meanwhile, the multi-objective optimization approaches are insufficient, especially in triple objectives. Therefore, developing an alternative solution approach is necessary to solve the DRCFJSP.

The wild horse optimizer (WHO) is a new population-based optimization algorithm developed by Naruei and Keynia [29]. The WHO imitates the social lift behavior of wild horses, especially for the decency behavior of the horse. Compared to the existing algorithms, such as GA, particle swarm optimization, salp swarm algorithm, etc., the WHO performs better in solving the CEC2019 test functions. Since then, the WHO attracted the attention of many scholars. Li *et al*. [30] suggested four strategies to improve the optimization capability of the WHO and performed a demonstration of the improvement on CEC2017 and CEC2021. Vasanthkumar *et al*. [31] introduced deep learning to enhance the WHO for the estimation of the state of charge in hybrid electric vehicles. Ali *et al*. [32] used the WHO to optimize the distributed generation to increase the reliability, stability, and security of the electrical power systems. Alphonse *et al*. [33] adopted the WHO to allocate electric vehicle charging stations and photovoltaic energy resources in a smart grid simultaneously. Milovanović *et al*. [34] studied the multi-objective energy management problem using the WHO. Based on the applications of the WHO, it is rare to apply the WHO to scheduling problems. In this paper, we tried to design an improved WHO algorithm to fill the research gap in discretizing the WHO for solving DRCFJSP.

3. Problem Description

The DRCFJSP is described as follows: n jobs $J = \{1,2,\cdots,n\}$ needs to be processed on m machines $M = \{1,2,\cdots,m\}$ accompanied by w workers $W = \{1,2,\cdots,w\}$. Each job includes multiple operations, and the processing sequence of the operations must be followed. Each machine can only perform one operation at the same time. The operations can be processed on one of the available machines, and workers with the ability to operate the corresponding machines are assigned according to the process requirements. The worker number w should be less than or equal to the machine number m, and the processing time is determined by the operation efficiency of the selected workers and the capacity of the assigned machine. Each worker can operate more than one machine, and the machining operation efficiency of each worker is different.

To evaluate the effectiveness of the scheduling scheme in terms of production efficiency and machine utilization, we have selected three objectives for optimization, which include minimizing the makespan, maximal machine workload, and total machine workload. Among these goals, the makespan is crucial in determining production efficiency as it directly relates to the completion time of each job. The maximal machine workload identifies the bottleneck machine, and reducing it can help improve the workshop's production efficiency. Additionally, the total machine workload is closely related to machine idle losses and energy consumption. The scheduling process aims to provide an optimal resource allocation scheme and process sequence while adhering to process constraints and dual resource capacity constraints. Achieving the optimal schedule involves minimizing the aforementioned three objectives.

Table 1 Notation

J	set of jobs, $J = \{1,2,\cdots,n\}$
M	set of machines, $M = \{1,2,\cdots,m\}$
W	set of workers, $W = \{1,2,\cdots,w\}$
W_k	set of eligible workers that can operate machine k, $W_k \subseteq W$
J_i	set of operations of job i $(i \in J)$, $J_i = \{1,2,\cdots,n_i\}$
O_{ij}	the j-th operation of job i $(i \in J, j \in J_i)$
C_i	the complete time of job i $(i \in J)$
s_{ij}	the starting time of the operation O_{ij}
c_{ij}	the completion time of the operation O_{ij}
p_{ijkr}	the processing time of operation O_{ij} is processed on machine k by worker r
Δ	a big number
x_{ijkr}	binary variable: 1, if the operation O_{ij} is processed on machine k by worker r; 0, otherwise
y_{ghijk}	binary variable: 1, if both operation O_{ij} and operation O_{gh} are processed on machine k, and O_{gh} is earlier than O_{ij}; 0, otherwise
z_{ghijr}	binary variable: 1, if both operation O_{ij} and operation O_{gh} are processed by worker r, and O_{gh} is earlier than O_{ij}; 0, otherwise
$T_{w_l}^S$	the idle start time for worker w_l
$T_{w_l}^E$	the idle end time for worker w_l

The notation is defined in Table 1. The mixed-integer programming model of the DRCFJSP with the triple objectives is given as follows.

$$\min f_1 = \max_{i \in N}\{C_i\} \tag{1}$$

$$\min f_2 = \max_{k \in M}\{\textstyle\sum_{r \in W_k} \sum_{i \in N} \sum_{j \in J_i} p_{ijkr} x_{ijkr}\} \tag{2}$$

$$\min f_3 = \textstyle\sum_{k \in M} \sum_{r \in W_k} \sum_{i \in N} \sum_{j \in J_i} p_{ijkr} x_{ijkr} \tag{3}$$

s.t.

$$\textstyle\sum_{k \in M} \sum_{r \in W_k} x_{ijkr} = 1 \;\; \forall j \in J_i, i \in N \tag{4}$$

$$c_{ij} \geq s_{ij} + \textstyle\sum_{k \in M} \sum_{r \in W_k} (p_{ijkr} x_{ijkr}) \; \forall j \in J_i, i \in N \tag{5}$$

$$c_{ij} \leq s_{i(j+1)} \; \forall j \in J_i \backslash \{n_i\}, i \in N \tag{6}$$

$$y_{ghijk} + y_{ijghk} \leq 1 \; \forall h \in J_g, j \in J_i, g, i \in N, k \in M \tag{7}$$

$$z_{ghijr} + z_{ijghr} \leq 1 \; \forall \; h \in J_g, j \in J_i, g, i \in N, r \in W \tag{8}$$

$$c_{gh} \leq s_{ij} + \triangle\,(3 - \textstyle\sum_{r \in W_k} x_{ghkr} - \sum_{r \in W_k} x_{ijkr} - y_{ghijk}) \forall h \in J_g, j \in J_i, k \in M, g, i \in N \tag{9}$$

$$c_{gh} \leq s_{ij} + \triangle\,(3 - \textstyle\sum_{k \in M} x_{ghkr} - \sum_{k \in M} x_{ijkr} - z_{ghijr}) \forall h \in J_g, j \in J_i, r \in W, g, i \in N \tag{10}$$

$$C_{max} \geq c_{ij} \; \forall j \in J_i, i \in N \tag{11}$$

$$x_{ijkr}, y_{ghijk}, z_{ghijk} \in \{0,1\} \forall h \in J_g, j \in J_i, r \in W_k, k \in M, g, i \in N \tag{12}$$

The objective functions consist of three parts. Eq. 1 minimizes the maximum makespan, Eq. 2 minimizes the maximum machine workload, and Eq. 3 minimizes the overall machine load. Constraint Eq. 4 ensures that each operation of a job can only be allocated to a single worker and performed simultaneously by a single piece of machine. The processing completion time of each operation is the sum of its earliest starting time and its processing time, as shown by Constraint Eq. 5. There is a processing sequence restriction between two adjacent job operations, according to constraint Eq. 6. The conflicting sequence of operations allocated to a machine is avoided by constraint Eq. 7. The conflicting sequence of operations allocated to a worker is avoided by constraint Eq. 8. Constraints Eqs. 9 to 10 establish the respective association between x, y, and z. Makespan is specified by constraint Eq. 11. The domains of three decision variables are set by constraint Eq. 12.

4. The improved wild horse optimization algorithm

Due to the intractable nature of the problem under study and the necessity of finding multiple Pareto optimal solutions, an improved multi-objective wild horse optimization algorithm is proposed. The wild horse optimizer (WHO) is a new meta-heuristic algorithm for solving continuous optimization problems [29]. This population-based optimization algorithm imitates the behavior of non-territorial wild horses to find the optimal solution in the solution space. Specifically, group behaviors, grazing, mating, dominance, and leadership are utilized to design search operators. As with other optimization algorithms, the WHO starts with an initial random population with N_{pop} members. Moreover, the G ($G = \lceil N_{pop} \times PS \rceil$, where PS is the percentage of stallions in the population) leaders (stallions) are selected based on their fitness values to determine the groups, and the remaining members ($N_{pop} - G$) are divided equally among these groups. The stallion can be regarded as the center of the grazing area, and the group members are searched around the center. When the position of the stallion is in a different direction, its members are attracted by using the position updating equation. Compared to other animals, the behavior of separating foals from the group and mating them prevents the father from mating with the daughter or siblings. In the WHO, the crossover operator is used to simulate the behavior of the departure and mating of horses. According to the horse mating behavior, the prematurity of the population can be prevented. The group leader must lead the group to a suitable area. The water hole can be taken as this suitable area. And the leaders must compete for this water hole so that the dominating group can use this water hole and other groups are not allowed to use the water hole. The WHO simulates this behavior to update the stallion's location and expedite the search process. In the initial iteration, the group's leaders are selected at random. In later iterations of the WHO, leaders are chosen based on their fitness. If one member of the group is superior to the group leaders, their positions will be switched. If the termination criteria are met, the algorithm is stopped, and the best solution is output. Otherwise, the algorithm is iterated.

The standard WHO algorithm was created to solve continuous optimization problems. Here, discrete optimization is required for the investigated DRCFJSP. The section introduces the specifics of the improved WHO algorithm. The enhanced components are designed based on the peculiarities of our problem.

4.1 Encoding and decoding scheme

In our algorithm, the encoding scheme adopts a three-vector string [21] to express the operation sequence (OS), machine assignment (MA), and worker assignment (WA) in a solution, as shown in Fig. 1. The first vector OS is the order of the jobs' operations. The number of occurrences for a specific job index corresponds to its operation number. The second vector MA and the third vector WA determine the corresponding suitable machine and worker for the specific operation of the job in OS. The lengths of the three vectors are $\sum_{i \in N} n_i$.

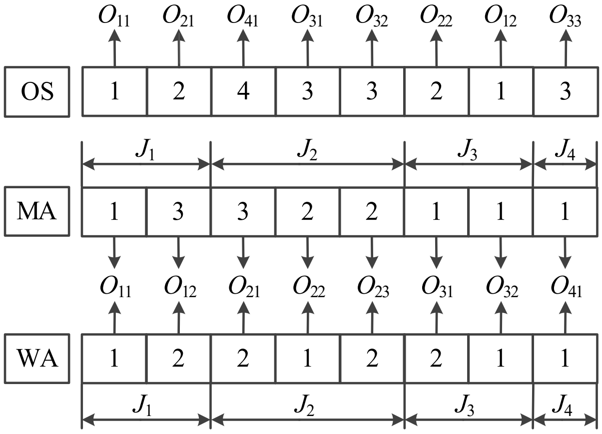


Fig. 1 Solution encoding scheme

Based on the above representation, the corresponding schedule can be obtained via the following five steps of the decoding procedure.

1) Take the operation O_{ij} from the OS one by one from left to right and find the corresponding machine k and worker r;
2) Based on the processing time p_{ijkr} of operation O_{ij} handling by machine k and worker r, the free interval $[T_k^S, T_k^E]$ of machine k and the free interval $[T_r^S, T_r^E]$ of worker r can be determined, in which T^S and T^E are the starting time and the ending time of the free interval, respectively.
3) Calculate the starting time T_{ij}^S of the operation O_{ij} via the Eq. 13;

$$T_{ij}^{s} = max(T_{i(j-1)}^{E}, T_k^S, T_r^S) \tag{13}$$

4) Evaluate whether the operation O_{ij} can be processed earlier. If $T_{ij}^S + P_{ijkr} \leq min(T_k^E, T_r^E)$, the operation O_{ij} can be inserted into the available free interval forward. If the operation O_{ij} can not be inserted into the existing free interval, it is appended to the processing list of the operations assigned to machine k. And the starting time of operation O_{ij} is the maximum between the ending time of the last operation assigned to machine k and the ending time of the last operation assigned to worker r;
5) Repeat the above steps 1-4 until all operations of the jobs have been processed.

4.2 Population initialization

To improve the WHO algorithm's population initialization, multiple groups of individuals should be initialized, in line with the algorithm's design philosophy. The exchange of information between these groups can expedite the search procedure, but the quality of initial solutions can also enhance the algorithm's ability to explore and exploit. To generate the OS, our algorithm selects a processing operation randomly from the list of candidate operations while maintaining the precedence relationship between the job operations. The MA and WA are then generated by assigning the most efficient machine and worker, respectively.

4.3 Pareto ranking and crowding distance

The fast-non-dominated sorting approach proposed in non-dominated sorting genetic algorithm II (NSGA-II) [35] is adopted to construct the set of Pareto optimal solutions. The sorting approach distinguishes all nondominated solutions based on the individuals' objective values and directs the population toward the set of Pareto optimal solutions. We consider two entities for each individual: 1) $n(i)$ represents the number of the solutions that dominate the individual i, and 2) $S(i)$, a set of individuals that the individual i dominates. The detailed computation process for the two entities can be found in reference [35]. Individuals with rank one belong to the non-dominated set which is also called the Pareto-optimal set. Once the non-dominated sorting is finished, a new population of size N_{pop} is formed with individuals of different non-dominated fronts.

After the Pareto ranking of the individuals, the priority relationship of the individuals located on the same front should be analyzed. In our improved WHO algorithm, individual sorting is used to guide the updating of the remaining individuals with the information of the best one. And the crowding distance is used to sort the individuals of the same front. The crowding distance of an individual from the nondominated solution set is the sum of the differences in the objective values of its two adjacent individuals. Here, we set the boundary value of the crowding distance to infinity for selecting other individuals at the next iteration. Before calculating the crowding distance, the individuals of various Pareto fronts are sorted based on their objective values. The following Eq. 14 is used to assign the crowding distance D_i to the individual i.

$$n_d = \sum_{h}^{H} \frac{f_h(i+1) - f_h(i-1)}{f_h^{\max} - f_h^{\min}} \tag{14}$$

In the above equation, H is the number of objective values at the same front, f_h^{max} is the maximum value of the objective function f_h, f_h^{min} is the minimum value of the objective function f_h, $f_h(i+1)$and $f_h(i-1)$ are the objective values of the two most adjacent individuals for i at the sorting sequence, respectively.

4.4 Position updating for group members

In the original WHO introduced in reference [29], the positions of group members are updated based on the grazing and mating behavior of horses. In our algorithm, the stallion of the group and the stallions of other groups guide the updating of the position of the individuals. Once the population has completed the Pareto ranking, and the calculation of crowding distance, the member with the greatest crowding distance is regarded as the group's stallion (the group leader). Eq. 15 motivates the group members to move and search around the leader of the same group or the leader of other groups.

$$X'_{i,j,(t+1)} = X_{i,j,t} \oplus X_{i,best,t} \oplus X_{k,best,t} \tag{15}$$

In the above equation, t is the iteration index, $X_{i,j,t}$ is the current position of the individual j of the group i at iteration t, $X_{i,best,t}$ is the position of the stallion of group i; $X_{k,best,t}$ is the position of the stallion of the group k ($k \neq i$). In particular, the group k is randomly selected from the remaining groups except for group i. Moreover, $\oplus$ is the crossover operator introduced later. Eq. 15 means the position of the group member should be updated by its group leader or another group leader. The grazing behavior and horse mating behavior can be embodied in Eq. 15. In the subsequent section, we introduce the crossover operator using Eq. 15.

Crossover operator

To improve the efficiency and effectiveness of the proposed algorithm, we implemented two crossover operators to obtain new group member positions. An improved precedence operation crossover (IPOX) is used to create the new sequence for the OS vector. The IPOX is not only able to retain excellent solutions for the subsequent iteration, but it can also effectively reduce the generation of solutions that are not feasible. The procedures of IPOX for OS are described in detail below. Each group member is paired with a random stallion from another group to form a crossover pair. PO1 and PO2 represent the parents. Let CO1 and CO2 represent their offspring. Some jobs are randomly selected to form a set S, where $|S| \leq n-1$. Then, all jobs are split into two sets S and $N-S$. The positions of the selected jobs S of parent PO1 are copied to CO1. The job of the set $N-S$ is taken from the PO2 in sequence and inserted into the remaining position of the CO1. The same procedure is executed for PO2 and CO2. An example of the IPOX operator is shown in Fig. 2.

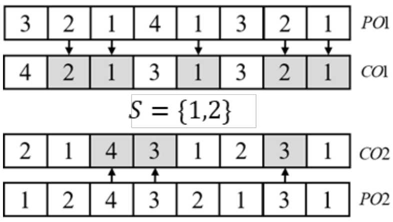


Fig. 2 IPOX for OS

Moreover, the multi-point crossover (MPX) is adopted to generate the new vectors for MA and WA, and the MPX randomly selects some points in the parent individuals for crossover. First, a random set R of 0-1 numbers is generated. In particular, $|R| = \sum_{i \in N} n_i$. The elements of R determine whether the crossover operator should be implemented. When the element is 1, the crossover should be implemented. Let PM1/PW1 and PM2/PW2 denote the parent individuals for MA and WA. Let CM1/CW1 and CM2/CW2 represent the offspring individuals for MA and WA. When the element of R is 1, the elements of two-parent individuals are exchanged to form the offspring individuals. Otherwise, the elements of two-parent individuals are copied into the offspring individuals. When the offspring individuals are infeasible, the parent individuals are kept. An example of an MPX operator is presented in Fig. 3.

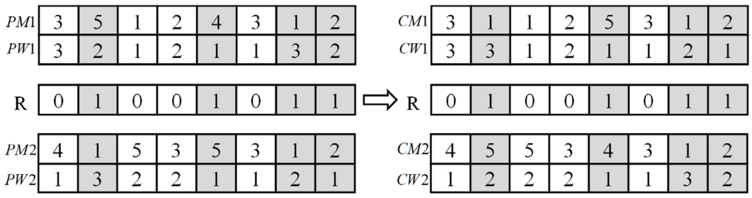
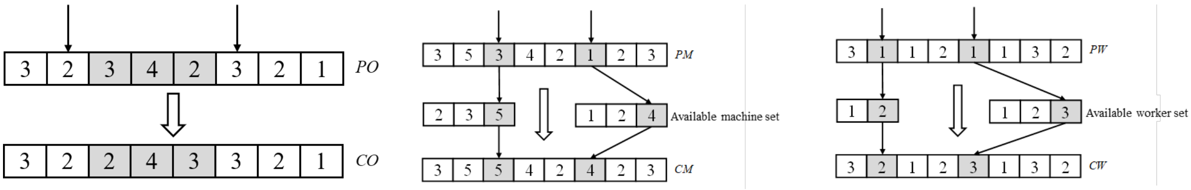


Fig. 3 MPX operator for MA and WA

Mutation operator

The mutation operator working on the elite individuals is turned to lasting diversity during the evolution process. This paper adopts two mutation operators for a three-level encoding scheme. The OS string uses the two-point reversing operator to generate the mutated individual. That is to say, two elements are selected randomly from the OS string, and all operations between the selected two elements are reversed. An example is shown in Fig. 4(a), where PO is the parent OS string, and CO is the chromosome OS string. The two-point random mutation operator is used for the MA and WA strings. In particular, one element is randomly selected from the MA string. And a different component is chosen from the set of available machines for the corresponding job operation. The above procedure is repeated again for another randomly chosen MA element. Moreover, the mutation operator of the WA is the same as the MA. Only the set of the available worker should be used. As shown in Fig. 4(b) and Fig. 4(c), the new MA string and the new WA string are generated based on the available candidate set.



(a) OS (b) MA (c) WA

Fig. 4 Mutation operator

4.5 Local search operator

The primary purpose of the local search strategy is to prevent premature convergence and ensure the global search ability of the proposed algorithm. To achieve this, a local search method based on non-dominated sorting is designed. While the introduction of a local search strategy may increase the algorithm's run time, this paper balances run time and search effect by sorting the individuals of each non-dominated solution set by crowding distance, selecting the first 50 % of individuals to construct a search set, and performing a local search on this set. The neighborhood is perturbed by modifying the processes on the critical path.

The neighborhood structures are designed based on the following rules: 1) The two key operations at the end of the first key block on the critical path are exchanged to satisfy process requirements while keeping the machine and worker indexes of the operation unchanged; 2) If the key block is at the end of the critical path, the two key operations at the beginning of the block are exchanged according to the process precedence relationship, without modifying the machine and worker indexes; 3) For other key block operations, only the key operations adjacent to the block head and block tail are exchanged, and operations in the remaining key blocks are not perturbed. No perturbation is performed if there is only one critical operation in the critical block. 4) To avoid generating infeasible solutions, perturbation is not performed if the exchange of operations belongs to the same workpiece. The local search procedure for our multi-objective problem is outlined as follows:

Step1: Construct the neighborhood solution set $NO = \emptyset$ and the comparison solution set $CP = \emptyset$;

Step2: For each individual s_i of search solution set S, all neighborhood solutions are generated by the designed neighborhood structure；

Step3: Evaluate the dominant relationship between the individual s_i and its neighborhood solution s_i'. If $s_i \prec s_i'$, the neighborhood solution s_i' is abandoned; if s_i and s_i' are irrelated, the neighborhood solution s_i' is added into the set NO. If $s_i' \prec s_i$, the current solution s_i is replaced by s_i', the neighborhood solution s_i' is added into the set CP, the neighborhood solution set NO is set to empty, and the search of the individual s_i is stopped;

Step4: Once all neighborhood solutions of the individual s_i have been searching, if a non-dominated solution of s_i is found, s_i is put into the set NO. And the set NO and the set CP are merged. The set NO is set to empty, and the local search starts to implement the above search process for the next individual s_{i+1};

Step5: After searching all individuals of the set S, all individuals of the set CP are nondominated sorted. The non-dominated solutions with large crowding distances are selected to replace the individuals of the set S.

The local search strategy proposed in this paper is based on non-dominated sorting for individual selection, which can effectively select the best solution in the population so that the algorithm can be accelerated toward the convergence to the Pareto frontier.

4.6 Framework of the improved WHO algorithm

The procedure of the improved WHO algorithm solving the DRCFJSP is depicted as follows. Firstly, the algorithm parameters are determined and the instance parameters are input into the algorithm. A population with N_{pop} individuals is generated. The three-level encoding scheme and the greedy insert decoding scheme are used to obtain the schedule with triple objectives. The population based on the fast non-dominated sorting and crowding distance is evaluated. Then, the best and worst solutions are selected to guide the individual updating. A pair of solutions is randomly selected out of the population, and the crossover operation is executed. The crossover process is repeated until every individual is paired with the other population individual with the greatest crowding distance. The individuals from various groups are selected to execute the mutation operation. The above process is repeated until the number of individuals meets the predetermined size. After that, the local search is performed on the first 50 % of individuals based on the crowding distance of the nondominated individuals at the same level. The above process is repeated until the number of individuals meets the predetermined size. Subsequently, the elitist archive with the first Pareto front solutions of the new population is updated. Once the termination criteria are met, output the best individual.

5. Experimental results and analysis

The proposed IWHO algorithm is implemented in Python. The code is run on a PC with an Intel Core i5-6400 CPU (2.70GHz), 8GB RAM, and a 64-bit Windows 10 operating system. For evaluating the performance of the improved IWHO algorithm in solving the multi-objective DRCFJSP problem, the standard NSGA-II method is used as a benchmark here. The two algorithms are run ten times for each test instance to eliminate the effect of the random search.

5.1 Instance generation

Two groups of test instances were generated using the method proposed in reference [17] based on the BRdata and DPpaulli. The first group consists of ten instances with MK01-10[36] and the second group includes eight instances with DP01-08 [37]. Except for the processing information of the jobs, we added the data of machines and workers into the two groups of instances, as shown in Table 2. In particular, the processing time of each operation for all jobs is randomly picked from the interval $[p_{ijk}, p_{ijk} + \delta_{ij}]$, where p_{ijk} is the processing time of the job operation in the original benchmark instance, and δ_{ij} is a random number from the interval [2, 8].

Table 2 Machine-worker data for test instances

Instances	w	The set of eligible workers operating machine k
DMK1-2	4	$W_1 = \{1,3\}, W_2 = \{2,4\}, W_3 = \{1,4\}, W_4 = \{2,3,4\}, W_5 = \{1,2\}, W_6 = \{3\}$
DMK3-4; DP7-8	6	$W_1 = \{1,3\}, W_2 = \{2,4\}, W_3 = \{4\}, W_4 = \{2,3\}, W_5 = \{1,6\}, W_6 = \{3,4,5\}, W_7 = \{4,5\}, W_8 = \{5,6\}$
DMK5	3	$W_1 = \{1,3\}, W_2 = \{2,3\}, W_3 = \{1,3\}, W_4 = \{1,2\}$
DMK6; DMK10	8	$W_1 = \{1,8\}, W_2 = \{2,4\}, W_3 = \{3,8\}, W_4 = \{3,7\}, W_5 = \{6\}, W_6 = \{5\}, W_7 = \{2,5\}, W_8 = \{1,5,6\}, W_9 = \{4,7\}, W_{10} = \{1,6,8\}, W_{11} = \{2,3\}, W_{12} = \{4\}, W_{13} = \{4\}, W_{14} = \{7,8\}, W_{15} = \{5,7\}$
DMK7; DP1-6	4	$W_1 = \{1,4\}, W_2 = \{2,4\}, W_3 = \{1,3\}, W_4 = \{2,3\}, W_5 = \{1,4\}$
DMK8, DMK9	6	$W_1 = \{1,4\}, W_2 = \{2,6\}, W_3 = \{1,3\}, W_4 = \{2,3,6\}, W_5 = \{1,5\}, W_6 = \{5\}, W_7 = \{4,5\}, W_8 = \{3,6\}, W_9 = \{2,4\}, W_{10} = \{3,6\}$

5.2 Parameter setting

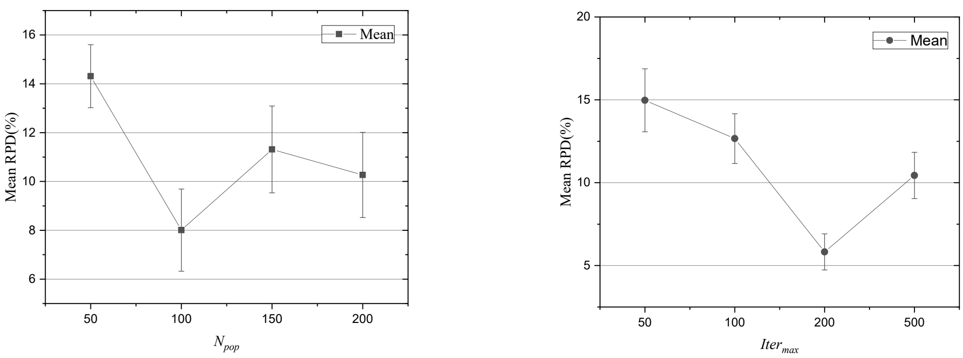
The IWHO algorithm does not need to tune the specific parameters involved in other population-based algorithms. Here we only considered two parameters, including population size N_{pop} and number of iterations $iter_{max}$. The candidate values of the two parameters are listed as follows. $N_{pop} \in \{50,100,150,200\}$, $iter_{max} \in \{50,100,200,500\}$.

Then, we had 16 combinations for the two parameters to tune. The IWHO solved the first four instances of Brdata with 16 combinations. The performance of each combination is measured by the relative percentage deviation (RPD) defined by $RPD(\%) = 100 \times (f_1(i) - f_1^*)/f_1^*$, where $f_1(i)$ is the makespan delivered by the IWHO with the *i*th parameter value and f_1^* is the best value obtained by the IWHO. The two-way variance analysis (ANOVA) test examined the computational results. The ANOVA results are shown in Table 3. The analysis indicates that parameters $iter_{max}$ and N_{pop} are significantly different. However, the interaction between the two parameters is not statistically significant due to its high *P*-value. That means it is unnecessary to consider the interaction in further analysis.

Furthermore, the two parameters are further analyzed by a multi-compared method. The means plot and 95 % confidence level Tukey's Honestly Significant Difference (HSD) intervals for these four candidate values of N_{pop} and $iter_{max}$ are described in Fig. 5(a) and Fig. 5(b), respectively. As can be seen from Fig. 5(a), Tukey's HSD interval with $N_{pop} = 100$ has a better mean RPD value. When considering the running time, we have chosen 100 as the best value of the parameter N_{pop}. The Tukey's HSD interval with $iter_{max}$=200 is significantly lower than that with the other three values. Therefore $iter_{max} = 200$ gives the best result for our algorithm. Therefore, we set the two parameters $N_{pop} = 100$ and $iter_{max} = 200$ in the following computations.

Table 3 ANOVA results for the experiment on tuning the parameters of IWHO

Source	DF	Sum of Squares	Mean Square	*F* Value	*P* Value
N_{pop}	3	329.008	109.669	3.042	0.038
$iter_{max}$	3	729.849	243.283	6.748	0.000
Interaction	9	104.434	11.603	0.321	0.963
Model	15	1163.293	77.552	2.151	0.023
Error	48	1730.395	36.049		
Corrected Total	63	2893.688			



(a) N_{pop} (b) $iter_{max}$

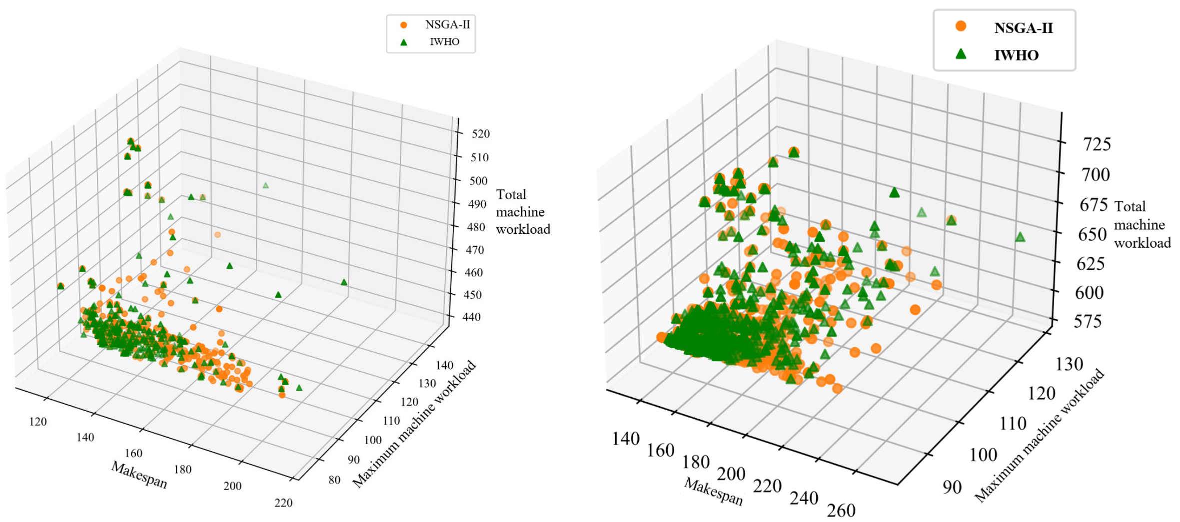
Fig. 5 Means plot and 99 % confidence level Tukey's HSD intervals for population size N_{pop} and $iter_{max}$.

5.3 Result analysis

To evaluate the effectiveness of the algorithm proposed in this paper, we compare the IWHO algorithm with the NSGA-II algorithm for solving the DRCFJSP. The computational results consist of three objectives: minimizing the maximum makespan, the maximal machine workload, and the total machine workload. Table 4 displays the computational results of the IWHO and NSGA-II algorithms for these three objectives, respectively. Moreover, the best objective values are highlighted. It can be observed that the solutions provided by our IWHO algorithm dominate those obtained by NSGA-II for most instances, indicating that the three objectives of the solutions given by IWHO are superior to those of NSGA-II. Upon analyzing the individual objectives, IWHO outperforms NSGA-II in terms of makespan. For total machine workload, 13 out of 18 instances achieved better solutions using IWHO, while for maximal machine workload, both algorithms performed comparably. The running time of the two algorithms is nearly identical. As a multi-objective problem, DRCFJSP requires balancing the three objectives. Our IWHO algorithm performs exceptionally well in solving the studied multi-objective scheduling problem.

Table 4 Computational Results for the IWHO algorithm

Instance	NSGA- II							IWHO			
	n	m	w	f_1	f_2	f_3	CPU(s)	f_1	f_2	f_3	CPU(s)
DMK01	10	6	4	76	52	240	1316	**71**	**52**	**224**	**1175**
DMK02	10	6	4	69	47	208	1586	**60**	**39**	**208**	**1211**
DMK03	15	8	6	310	220	1087	1964	**272**	**219**	**1043**	**1662**
DMK04	15	8	6	144	85	468	1531	**121**	87	**448**	**1228**
DMK05	15	4	3	370	210	814	1576	**332**	**210**	819	**1546**
DMK06	10	15	8	136	88	576	1672	**135**	97	**566**	1737
DMK07	20	5	4	304	185	792	1494	**268**	191	793	**1264**
DMK08	20	10	6	798	613	2833	1912	**744**	**604**	2882	**1903**
DMK09	20	10	6	661	348	2616	1878	**624**	**347**	2680	2045
DMK10	20	15	8	486	266	2258	2167	**462**	**261**	**2256**	**2066**
DP01	10	5	4	4667	2556	11368	1785	**3962**	2574	11418	**1644**
DP02	10	5	4	4586	2285	11361	1831	**3967**	2329	**11352**	**1774**
DP03	10	5	4	4157	2326	11311	1907	**4116**	**2307**	11360	**1836**
DP04	10	5	4	4693	2576	11354	1819	**3888**	2656	**11339**	1867
DP05	10	5	4	4140	2284	11251	1709	**3902**	2384	**11250**	1892
DP06	10	5	4	4418	2286	11132	1930	**3948**	**2269**	11214	**1920**
DP07	15	8	6	5247	2337	16863	2089	**4406**	2387	16925	2196
DP08	15	8	6	5209	2138	16696	2458	**4366**	2201	16748	**2006**



(a)DMK04 (a)DMK06

Fig. 6 Pareto front for DMK04 and DMK06 solutions delivered by NSGA-II and IWHO

To provide a visual representation of the performance of the two algorithms, a 3D scatter plot was used to plot the Pareto optimal fronts generated by the solutions for DMK04 and DMK06, as depicted in Fig. 6. The scatter plot clearly reveals the trade-offs between the three objectives being considered, indicating a trade-off between makespan and total machine workload, as well as between total machine workload and maximum machine workload. Additionally, there is a positive correlation between makespan and maximum machine workload. The IWHO algorithm Pareto fronts are distributed evenly and are positioned closer to the bottom of the graph, indicating its superior performance over NSGA-II.

Furthermore, the performance of the proposed IWHO algorithm is analyzed when changing the scale of the problem instances. The gap of the IWHO algorithm compared with the NSGA-II is calculated, such as gap $= \frac{f(\text{NSGA-II}) - f(\text{IWHO})}{f(\text{NSGA-II})} \times 100\%$. In particular, the gap value is larger, the proposed IWHO is better. When the gap value is greater than 0, the IWHO performs better than the NSGA-II. The size of the test instances is increased for the instances DMK01-DMK10. The gap curve of the proposed IWHO algorithm related to the NSGA-II is given in Fig. 7. It can be observed that the proposed IWHO algorithm tends to optimize the first objective (makespan) more, and there is no significant deterioration when the scale of the problem increases. That means the IWHO algorithm still performs better than NSGA-II, even increasing the problem complexity.

Fig. 7 The gap curve of the three objectives given by IWHO compared with NSGA-II

6. Real case study

To evaluate the practical applicability of the proposed algorithm, we conducted a case study at a railway rolling stock corporation located in Qingdao. The assembly process of the buffers and chain coupler for the high-speed rail was selected as the subject of our study, specifically focusing on the automatic buffers and chain coupler module. This module consists of several components, including the suspension system, buffer crushing system, connecting system, electrical control system, pneumatic control system, and connecting ring, as depicted in Fig. 8.

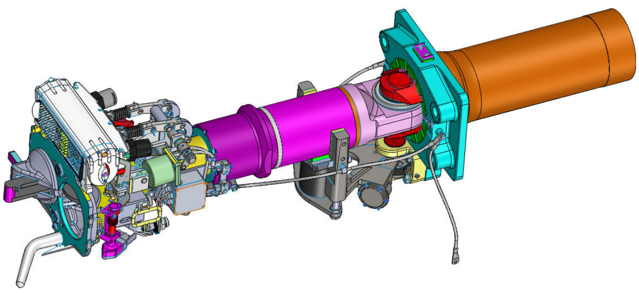


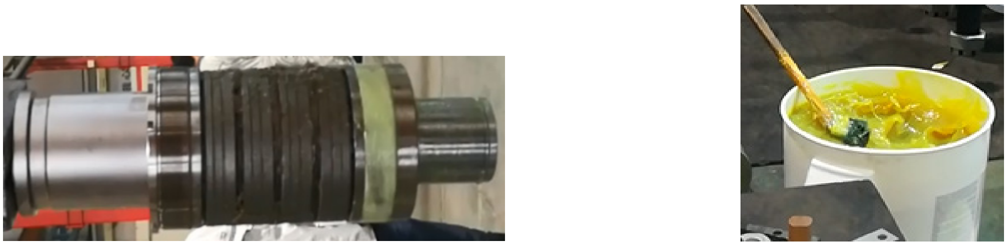
Fig. 8 Structure diagram of automatic buffers and chain coupler

The general assembly process of buffers and chain couplers is detailed as follows:

(1) To transfer the connecting system, a self-service crane hoists the subsequent connected system shown in Fig. 9 and places it on the spare parts table, where it awaits the following operation.

Fig. 9 Schematic diagram of the connecting system

(2) Lubrication of the shaft end of the buffer parts on Fig. 10(a) involves applying lubricating grease to the mating surface of the shaft end of the buffer parts using a brush, Fig. 10(b).



(a) Buffer parts (b) Grease

Fig. 10 Lubricating the shaft end of the buffer composition

(3) In preparation for composition, a press is used to drive two cylindrical pins into the pin holes of the lower connecting ring, and the guide plate is installed. The anti-loosening plate is tightened and torqued with hexagon head bolts to a torque of 70 Nm. Both ends of the anti-loosening plate are bent towards the hexagon head of the abutting bolt to prevent loosening. Once the lower connecting ring components are prepared, the inner ring surface is greased. The other elements of the connecting ring, including the upper half ring, connecting ring fastening bolts, nuts, anti-loosening plates, stop blocks, etc., are also prepared. Lubricating grease is applied to the inner surface of the upper half ring, and an appropriate amount of thread grease is applied to the threaded part of the bolt according to the working instructions. The whole process is shown in Fig. 11.

(a) Press (b) Lower connecting ring (c) Thread grease (d) Upper connecting ring

Fig. 11 Schematic of preparing the connection ring

(4) Installing the connection ring involves hoisting the connecting system to position the end of the rear crushing and centering buffer device and adjusting the position of the guide plate. Then, the connecting ring, stop block, bolt, anti-loosening plate, and nut are installed in sequence. The distance between the upper and lower rings is checked and adjusted to maintain the same spacing. Each nut is tightened three times in diagonal order, and torque is applied. Finally, the anti-loosening plate is blended until it is against the nut to prevent loosening.

(5) To install the cover plate, hexagon head bolts and anti-loosening plates are used to tighten and torque. An appropriate amount of thread grease is applied to the threaded parts of the bolts before fastening. After tightening the bolts, both ends of the anti-loosening plate are blended towards the side of the hexagonal head of the bolt to prevent loosening.
(6) The ground wire is installed, and the hexagon head bolts, anti-loose washers, and flat washers are connected and fastened.
(7) All tightened bolts with anti-loosening marks are scratched to signify their status.
(8) The grounding mark is pasted next to the terminal, and a layer of varnish is applied after pasting to protect it.
(9) Lubricating grease is applied to the circumference of the connecting bolts of the connecting ring and the mounting holes of the lower half ring.

The procedure outlined previously is identified by numbers (1) through (9). The assembly process for additional buffer and chain coupler types is comparable to the aforementioned process. However, the semi-permanent coupler that lacks a buffer does not necessitate the lubrication process for the shaft end (process 2). The workshop has 20 sets of equipment or consoles in nine groupings designated M1 through M20. The equipment accessible for each process varies, as indicated in Table 5.

Due to the variations in structure, weight, and process requirements of different coupler products, the time required for each process may vary. Additionally, the processing time for the same process step may vary due to differences in worker skills and proficiency. Table 6 presents the processing times of different processes for the automatic buffer and chain coupler used in high-speed rail for different workers. Please note that a blank value indicates that the worker assigned to that task is unable to perform the process.

The production data of the above automatic buffers and chain coupler was used in the DRCFJSP. Detailed test data can be obtained from the authors. The IWHO algorithm was used to solve the problem based on the best parameter combination determined by the ANOVA. The Gantt chart can be found in Fig. 12. Based on the comparison between the developed IWHO algorithm and the practical sequencing approach, and we found that the results delivered by our IWHO can achieve a 21.4 % improvement in the average three objective values.

Table 5 Available equipment for different assembly processes

Process	Available equipment
Process1	M1, M2
Process2	M3, M4, M5, M6
Process3	M7, M8, M9
Process4	M10, M11, M12
Process5	M10, M11, M12, M13, M14
Process6	M15, M16, M17, M18
Process7	M17, M18
Process8	M19, M20
Process9	M3, M4, M5, M6, M7

Table 6 Process timetable of the automatic buffers and chain coupler

Processes	Worker 1	Worker2	Worker3	Worker4	Worker5	Worker6	Worker7	Worker8
Process1	19		23					20
Process2	8	9	7	13	10	11	12	7
Process3	35	31			32	32		33
Process4			37	37	40	38	39	40
Process5					48	46	47	48
Process6		53					47	50
Process7	11	11					9	7
Process8			10	10	13	14		11
Process9	17	18	16	20	17	18	19	17

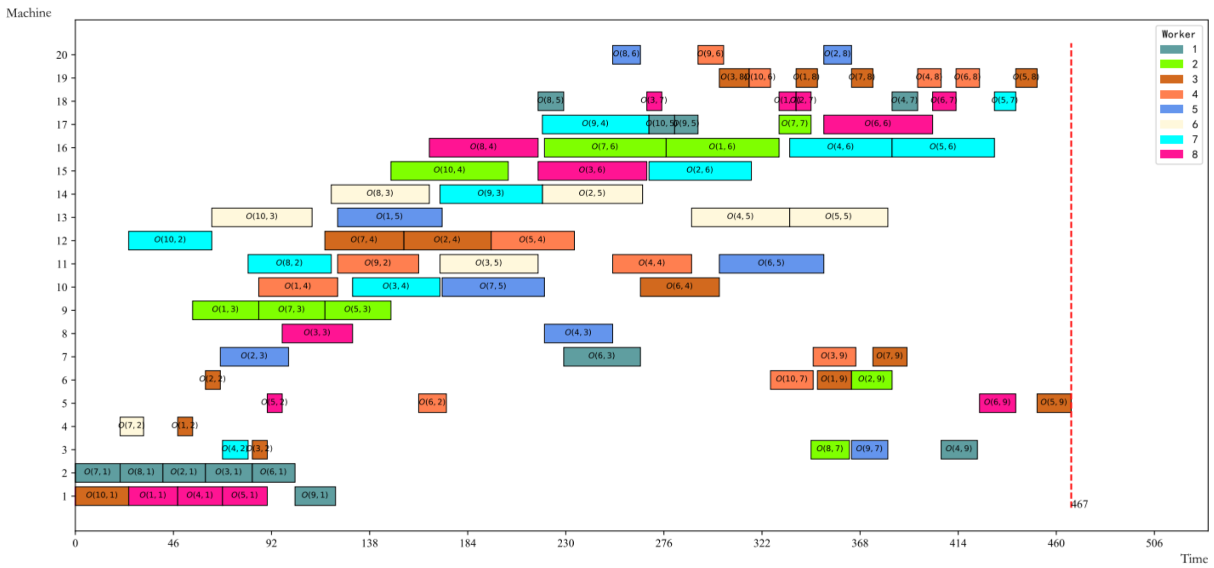


Fig. 12 Gantt chart for machines and workers in 10-job instance

7. Conclusions and future work

This paper proposes an IWHO algorithm designed to reduce makespan, maximum machine workload, and total machine workload for the DRCFJSP. The algorithm utilizes multi-group communication and three-level integer coding and implements a greedy insertion decoding technique to establish the initial schedule, taking into account the three optimization objectives. The algorithm also uses an elite preservation strategy to identify the best individuals in the population and incorporates the fast non-dominated sorting and crowding distance mechanism from the NSGA-II algorithm to sort the population and guide the WHO algorithm's members in updating their positions. To adhere to the characteristics of the WHO algorithm, the algorithm employs a crossover operator and a mutation operator to discretize the updating process. A local neighborhood search strategy centered on the critical path is introduced to prevent premature convergence and help avoid the local optimum. The computational test results demonstrate the effectiveness of the proposed algorithm for solving multi-objective DRSFJSP. Furthermore, a practical test scenario is designed to evaluate the performance of the suggested IWHO in assembling buffers and chain couplers. In the future, the study intends to explore dynamic problems in the workshop environment under resource constraints, such as rescheduling flexible job shops during emergencies such as order insertion, order cancellation, machine failure, or worker departure. Additionally, more effective improvement strategies, including simulated annealing, tabu search, etc., are encouraged to enhance the efficiency of the WHO algorithm.

References

[1] Çaliş, B. Bulkan, S. (2015). A research survey: Review of AI solution strategies of job shop scheduling problem, *Journal of Intelligent Manufacturing*, Vol. 26, No. 5, 961-973, doi: 10.1007/s10845-013-0837-8.

[2] Xie, J., Gao, L., Peng, K., Li, X., Li, H. (2019). Review on flexible job shop scheduling, *IET Collaborative Intelligent Manufacturing*, Vol. 1, No. 3, 67-77, doi: 10.1049/iet-cim.2018.0009.

[3] Amjad, M.K., Butt, S.I., Anjum, N., Chaudhry, I.A., Faping, Z., Khan, M. (2020). A layered genetic algorithm with iterative diversification for optimization of flexible job shop scheduling problems, *Advances in Production Engineering & Management*, Vol. 15, No. 4, 377-389, doi: 10.14743/apem2020.4.372.

[4] Lei, K., Guo, P., Zhao, W.C., Wang, Y., Qian, L., Meng, X., Tang, L. (2022). A multi-action deep reinforcement learning framework for flexible Job-shop scheduling problem, *Expert Systems with Applications*, Vol. 205, Article No. 117796, doi: 10.1016/j.eswa.2022.117796.

[5] Behnamian, J., Fatemi Ghomi, S.M.T. (2016). A survey of multi-factory scheduling, *Journal of Intelligent Manufacturing*, Vol. 27, No. 1, 231-249, doi: 10.1007/s10845-014-0890-y.

[6] Brucker, P., Schlie, R. (1990). Job-shop scheduling with multi-purpose machines, *Computing*, Vol. 45, 369-375, doi: 10.1007/BF02238804.

[7] Zhao, Z., Lin, P., Shen, L., Zhang, M., Huang, G.Q. (2020). IoT edge computing-enabled collaborative tracking system for manufacturing resources in industrial park, *Advanced Engineering Informatics*, Vol. 43, Article No. 101044, doi: 10.1016/j.aei.2020.101044.
[8] Xu, J., Xu, X., Xie, S.Q. (2011). Recent developments in Dual Resource Constrained (DRC) system research, *European Journal of Operational Research*, Vol. 215, No. 2, 309-318, doi: 10.1016/j.ejor.2011.03.004.
[9] Guo, D., Zhong, R.Y., Ling, S., Rong, Y., Huang, G.Q. (2020). A roadmap for Assembly 4.0: Self-configuration of fixed-position assembly islands under Graduation Intelligent Manufacturing System, *International Journal of Production Research*, Vol. 58, No. 15, 4631-4646, doi: 10.1080/00207543.2020.1762944.
[10] Xu, S., Hall, N.G. (2021). Fatigue, personnel scheduling and operations: Review and research opportunities, *European Journal of Operational Research*, Vol. 295, No. 3, 807-822, doi: 10.1016/j.ejor.2021.03.036.
[11] Grosse, E.H., Calzavara, M., Glock, C.H., Sgarbossa, F. (2017). Incorporating human factors into decision support models for production and logistics: Current state of research, *IFAC-PapersOnLine*, Vol. 50, No. 1, 6900-6905, doi: 10.1016/j.ifacol.2017.08.1214.
[12] Nourmohammadi, A., Fathi, M., Ng, A.H.C. (2022). Balancing and scheduling assembly lines with human-robot collaboration tasks, *Computers & Operations Research*, Vol. 140, Article No. 105674, doi: 10.1016/j.cor.2021.105674.
[13] Chaudhry, I.A., Khan, A.A. (2016). A research survey: Review of flexible job shop scheduling techniques, *International Transactions in Operational Research*, Vol. 23, No. 3, 551-591, doi: 10.1111/itor.12199.
[14] Li, H., Duan, J., Zhang, Q. (2021). Multi-objective integrated scheduling optimization of semi-combined marine crankshaft structure production workshop for green manufacturing, *Transactions of the Institute of Measurement and Control*, Vol. 43, No. 3, 579-596, doi: 10.1177/0142331220945917.
[15] Ren, W., Wen, J., Yan, Y., Hu, Y., Guan, Y., Li, J. (2021). Multi-objective optimisation for energy-aware flexible job-shop scheduling problem with assembly operations, *International Journal of Production Research*, Vol. 59, No. 23, 7216-7231, doi: 10.1080/00207543.2020.1836421.
[16] Ham, A. (2017). Flexible job shop scheduling problem for parallel batch processing machine with compatible job families, *Applied Mathematical Modelling*, Vol. 45, 551-562, doi: 10.1016/j.apm.2016.12.034.
[17] Lei, D., Guo, X. (2014). Variable neighbourhood search for dual-resource constrained flexible job shop scheduling, *International Journal of Production Research*, Vol. 52, No. 9, 2519-2529, doi: 10.1080/00207543.2013.849822.
[18] Lei, D., Guo, X. (2015). An effective neighborhood search for scheduling in dual-resource constrained interval job shop with environmental objective, *International Journal of Production Economics*, Vol. 159, 296-303, doi: 10.1016/j.ijpe.2014.07.026.
[19] Yazdani, M., Zandieh, M., Tavakkoli-Moghaddam, R., Jolai, F. (2015). Two meta-heuristic algorithms for the dual-resource constrained exible job-shop scheduling problem, *Scientia Iranica*, Vol. 22, No. 3, 1242-1257.
[20] Ahmadi, E., Zandieh, M., Farrokh, M., Emami, S.M. (2016). A multi objective optimization approach for flexible job shop scheduling problem under random machine breakdown by evolutionary algorithms, *Computers & Operations Research*, Vol. 73, 56-66, doi: 10.1016/j.cor.2016.03.009.
[21] Zheng, X.-L., Wang, L. (2016). A knowledge-guided fruit fly optimization algorithm for dual resource constrained flexible job-shop scheduling problem, *International Journal of Production Research*, Vol. 54, No. 18, 5554-5566, doi: 10.1080/00207543.2016.1170226.
[22] Xu, H., Bao, Z.R., Zhang, T. (2017). Solving dual flexible job-shop scheduling problem using a Bat Algorithm, *Advances in Production Engineering & Management*, Vol. 12, No. 1, 5-16, doi: 10.14743/apem2017.1.235.
[23] Ham, A. (2018). Scheduling of dual resource constrained lithography production: Using CP and MIP/CP, *IEEE Transactions on Semiconductor Manufacturing*, Vol. 31, No. 1, 52-61, doi: 10.1109/TSM.2017.2768899.
[24] Wu, R., Li, Y., Guo, S., Xu, W. (2018). Solving the dual-resource constrained flexible job shop scheduling problem with learning effect by a hybrid genetic algorithm, *Advances in Mechanical Engineering*, Vol. 10, No. 10, doi: 10.1177/1687814018804096.
[25] Wu, X., Peng, J., Xiao, X., Wu, S. (2021). An effective approach for the dual-resource flexible job shop scheduling problem considering loading and unloading, *Journal of Intelligent Manufacturing*, Vol. 32, No. 3, 707-728, doi: 10.1007/s10845-020-01697-5.
[26] Tan, W., Yuan, X., Wang, J., Zhang, X. (2021). A fatigue-conscious dual resource constrained flexible job shop scheduling problem by enhanced NSGA-II: An application from casting workshop, *Computers & Industrial Engineering*, Vol. 160, Article No. 107557, doi: 10.1016/j.cie.2021.107557.
[27] Andrade-Pineda, J.L., Canca, D., Gonzalez-R, P.L., Calle, M. (2020). Scheduling a dual-resource flexible job shop with makespan and due date-related criteria, *Annals of Operations Research*, Vol. 291, No. 1-2, 5-35, doi: 10.1007/s10479-019-03196-0.
[28] Vital-Soto, A., Baki, M.F., Azab, A. (2022). A multi-objective mathematical model and evolutionary algorithm for the dual-resource flexible job-shop scheduling problem with sequencing flexibility, *Flexible Services and Manufacturing Journal*, Vol. 35, 626-668, doi: 10.1007/s10696-022-09446-x.
[29] Naruei, I., Keynia, F. (2022). Wild horse optimizer: A new meta-heuristic algorithm for solving engineering optimization problems, *Engineering with Computers*, Vol. 38, Suppl. 4, 3025-3056, doi: 10.1007/s00366-021-01438-z.
[30] Li, Y., Yuan, Q., Han, M., Cui, R. (2022). Hybrid multi-strategy improved wild horse optimizer, *Advanced Intelligent Systems*, Vol. 4 , No. 10, Article No. 2200097, doi: 10.1002/aisy.202200097.
[31] Vasanthkumar, P., Revathi, A.R., Ramya Devi, G., Kavitha, R.J., Muniappan, A., Karthikeyan, C. (2022). Improved wild horse optimizer with deep learning enabled battery management system for internet of things based hybrid

electric vehicles, *Sustainable Energy Technologies and Assessments*, Vol. 52, Part C, Article No. 102281, doi: 10.1016/j.seta.2022.102281.

[32] Ali, M.H., Kamel, S., Hassan, M.H., Tostado-Véliz, M., Zawbaa, H.M. (2022). An improved wild horse optimization algorithm for reliability based optimal DG planning of radial distribution networks, *Energy Reports*, Vol. 8, 582-604, doi: 10.1016/j.egyr.2021.12.023.

[33] Alphonse, A.R.A., Raj, A.P.P.G., Arumugam, M. (2022). Simultaneously allocating electric vehicle charging stations (EVCS) and photovoltaic (PV) energy resources in smart grid considering uncertainties: A hybrid technique, *International Journal of Energy Research*, Vol. 46, No. 11, 14855-14876, doi: 10.1002/er.8187.

[34] Milovanović, M., Klimenta, D., Panić, M., Klimenta, J., Perović, B. (2022). An application of Wild Horse Optimizer to multi-objective energy management in a micro-grid, *Electrical Engineering*, Vol. 104, 4521-4541, doi: 10.1007/s00202-022-01636-y.

[35] Deb, K., Pratap, A., Agarwal, S., Meyarivan, T. (2002). A fast and elitist multi-objective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 2, 182-197, doi: 10.1109/4235.996017.

[36] Brandimarte, P. (1993). Routing and scheduling in a flexible job shop by tabu search, *Annals of Operations Research*, Vol. 41, No. 3, 157-183, doi: 10.1007/BF02023073.

[37] Dauzère-Pérès, S., Paulli, J. (1997). An integrated approach for modeling and solving the general multiprocessor job-shop scheduling problem using tabu search, *Annals of Operations Research*, Vol. 70, 281-306, doi: 10.1023/A:1018930406487.

APEM journal

Advances in Production Engineering & Management
Volume 18 | Number 3 | September 2023 | pp 288–302
https://doi.org/10.14743/apem2023.3.473

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

A feed direction cutting force prediction model and analysis for ceramic matrix composites C/SiC based on rotary ultrasonic profile milling

Amin, M.[a], **Rathore, M.F.**[b], **Ahmed, A.**[b], **Saleem, W.**[c,*], **Li, Q.**[d], **Israr, A.**[a]

[a]Institute of Space Technology, Islamabad, Pakistan
[b]School of Engineering, University of Jeddah, Jeddah, Kingdom of Saudi Arabia
[c]Technological University Dublin, Dublin, Ireland
[d]School of Mechanical Engineering & Automation, Beihang University, Beijing, P.R. China

A B S T R A C T

Ceramic matrix composites have immense applications in the aerospace, aircraft, and automobile industries. Belonging to this class, carbon-fiber reinforced ceramic matrix composites (C/SiC) are used for critical applications due to their superior properties. However, these materials have also stringent properties of heterogeneity, anisotropy, and varying thermal properties that affect machining quality and process efficiency. So, developing a cutting force prediction model and analyzing machining parameters is an essential need for the accurate machining of such materials. In this study, a mechanistic-based feed direction cutting force prediction model for rotary ultrasonic profile milling of C/SiC composites is developed and validated experimentally. The experimental and simulation results closely match each other. The mean error and standard deviation were recorded as 1.358 % and 6.003, respectively. The parametric sensitivity analysis showed that cutting force decreased with increased cutting speed, whereas it increased with increased feed rate and cutting depth. The proposed cutting force model for rotary ultrasonic profile milling of C/SiC composites is robust and can be applied to predict cutting forces and optimize the machining process parameters at the industry level.

A R T I C L E I N F O

Keywords:
Rotary ultrasonic profile milling;
Modeling;
Ceramic matrix composites C/SiC;
Brittle fracture;
Cutting force;
Machining process optimization

**Corresponding author:*
waqas.saleem@tudublin.ie
(Saleem, W.)

Article history:
Received 27 October 2023
Revised 5 November 2023
Accepted 7 November 2023

1. Introduction

Ceramic matrix composites have gained growing attention in aerospace, automobile, and high-tech industries. Carbon fiber-reinforced ceramic matrix composites (C/SiC) exhibit attractive properties, which makes them an ideal candidate for diverse applications. For example, C/SiCs have a stable coefficient of friction, excellent wear resistance, high-run-in performance, and exceptional thermal stability at high temperatures [1, 2]. The typical usage of such materials includes developing brake discs (used in F16 Fighter, Porsche GT2, and French TGV NG) and manufacture of critical structural parts (e.g., nose cone, guide vane, and wings) for new generation aerospace vehicles and hyper-sonic vehicles [3]. These materials are also applied for developing nose cones and nozzles of rocket engines due to their better temperature resistance and lightweight properties [4, 5]. However, due to high brittleness, anisotropy, and heterogeneity, the desired quality and efficiency of machined parts of such materials are challenging issues. High mechanical/ thermal loads and severe/ rapid tool wear were investigated [6]. So, machining difficult-to-cut materials with desired quality/ process efficiency was rigorous [7]. For efficient machining, cutting forces must be controlled within acceptable technical limits, as cutting force is the main index of the machining process [8, 9]. Due to excessive cutting forces, machining-induced defects, geometric/dimensional errors, and rapidly wear-off cutting tools were investigated [10]. Even though these materials developed near-to-net shapes, some machining processes are necessarily required to achieve final dimensions and surfaces [11]. With conventional machining processes like turning, drilling, milling, and grinding, issues related to quality and machining-induced defects are found in the machining of composite materials. Later, machining was found to improve with the invention of non-traditional machining processes such as electric discharge machining [38], ultrasonic machining, water jet machining, ultrasonic machining [39], vibration-assisted machining, rotary ultrasonic machining [31], electrolytic machining, chemical machining [12-14]. These machining technologies have shown better results for Ceramic matrix composites (CMC), Metal matrix composites (MMC), and Polymer matrix composites (PMC).

Rotary ultrasonic machining (RUM) was applied to drill glass material in 1966 [15]. Pei *et al*. [16] investigated rotary ultrasonic machining for milling of ceramics investigated. The hybrid machining process combines the material removal mechanism of diamond grinding process and ultrasonic machining. During this, a diamond abrasive core tool is ultrasonically vibrated in a normal direction with simultaneous spindle rotations. Prabhakar, 1992 worked on abrasive grit travels along its sinusoidal trajectory, causing hammering, abrasion, and extracting of workpiece material. The material removal occurs due to brittle fracture and the material flow plastically [16]. RUM is considered a rotary ultrasonic milling process if the abrasive core tool's feed direction is perpendicular to ultrasonic vibration's direction. However, when ultrasonic vibrations are applied in parallel to feed direction of abrasive core tool, the process is categorized as rotary ultrasonic drilling.

The published studies have shown that RUM demonstrates better machining for hard and brittle composite and ceramics. The cutting force is significantly reduced with RUM for ceramic matrix composites [16-18]. The improved surface integrity was investigated at the hole exit surface due to alteration of fiber fracture mechanism in rotary ultrasonic drilling of C/SiC composites [19]. Wang *et al*. [20] investigated that reduction in tearing size occurred by 30 % at hole exit with compound drill in rotary ultrasonic drilling of ceramic composite materials. The increase in ductile percentage was also investigated with increased spindle speed, while it decreased with increased vibration amplitude [21]. Hocheng *at al.* [22] analyzed improved machinability for rotary ultrasonic drilling of C/SiC composites. Li *at al.* [23] focused on advantages of RUM, considering cutting forces, material removal rate, and surface quality for CMC composites. Jiao *at al.* [24] studied the spindle speed and feed rate impact on cutting force. Bertsche *at al.* [25] found a significant decrease in cutting forces and tool wear for rotary ultrasonic slot milling of ceramic matrix composites. Ding *at al.* found a reduction in normal grinding force by 9-12 % and tangential grinding force by 9.7-19.4 %. The surface/sub-surface breakage decreased with ultrasonic grinding due to reduced grinding and ground surface roughness by 12 % compared to conventional grinding [26]. Yuan *at al.* [27] investigated the transition of ductile to brittle mode at a 4 μm depth of cut for rotary ultrasonic face milling of C/SiC composites.

In published literature, cutting force prediction models have been proposed to control cutting forces within acceptable limits. For instance, Yuan *at al.* [28] proposed a cutting force model for rotary ultrasonic face milling based on ductile mode for C/SiC composites. The cutting force models for rotary ultrasonic face milling of ceramics matrix composites (C/SiC) were established [29, 30]. Bertsche *at al.* [31] developed an analytical ultrasonic slot milling model of ceramic matrix composites.

The published studies have also reported the parametric sensitivity analysis and development of cutting force models for RUM of CMCs, like drilling, face, side, and slot milling. Other composite materials like glass (K9, BK7), PMC, and MMC have also been investigated for such machining processes. However, rotary ultrasonic machining for profile/ contour milling has rarely been reported for composite materials. Profile milling is widely used in machining of composite materials.

Keeping in view the challenges related to machining of C/SiC composites, there is an immense need for parametric investigation and development of a cutting force model for rotary ultrasonic profile milling of C/SiC composites to predict and control machining cutting forces to achieve better quality.

Novelty of the research

To the authors' knowledge, no study has been reported for rotary ultrasonic profile milling of composite materials for parametric analysis or cutting force model. The cutting force prediction model for rotary ultrasonic profile milling is reported in this research work. The presented research work is novel and provides new directions for machining composite materials.

This study develops a mechanistic-based model to predict feed direction cutting forces for rotary ultrasonic profile milling of Ceramic matrix composites-C/SiC. The model is developed by considering indentation fracture theory, brittle fracture, material removal mechanism, penetration trajectory, energy conservation theorem, and mathematical rules. The mathematical relationship of feed-cutting forces with parameters related to the machining process, workpiece material, and tool has been established. The cutting force prediction model is validated through data from experimental rotary ultrasonic profile milling of C/SiC composites.

The relationships between cutting force and machining parameters are investigated. This paper is organized into five sections. After the introduction, a mechanistic-based feed direction cutting force prediction model is developed in section 2. Section 3 explains experimental rotary ultrasonic profile milling for C/SiC composites. The results and discussion are covered in section 4. Finally, conclusions are presented in section 5.

2. Cutting force prediction model

This study applies rotary ultrasonic profile milling (RUPM) as the combination of ultrasonic vibration, grinding, and milling process, particularly with ultrasonic vibration perpendicular to the feed direction. During machining, the diamond abrasive core tool vibrates with ultrasonic frequency following a sinusoidal vibration path. The abrasive core tool's abrasive grits perform hammering, abrasion, and extraction (in sequence) in machining process. The material removal mechanism is based on indentation fracture theory. The rotary ultrasonic profile milling and trajectory of an abrasive grit are shown in Figs. 1(a), 1(b), respectively. The parameters and variables used in this study are given in Table 1. The following assumptions are made for the development of feed direction cutting force model: (a) diamond abrasive grits are rigid regular octahedron, (b) all diamond abrasive grits are of the same size, (c) material removal mode is a rigid brittle fracture.

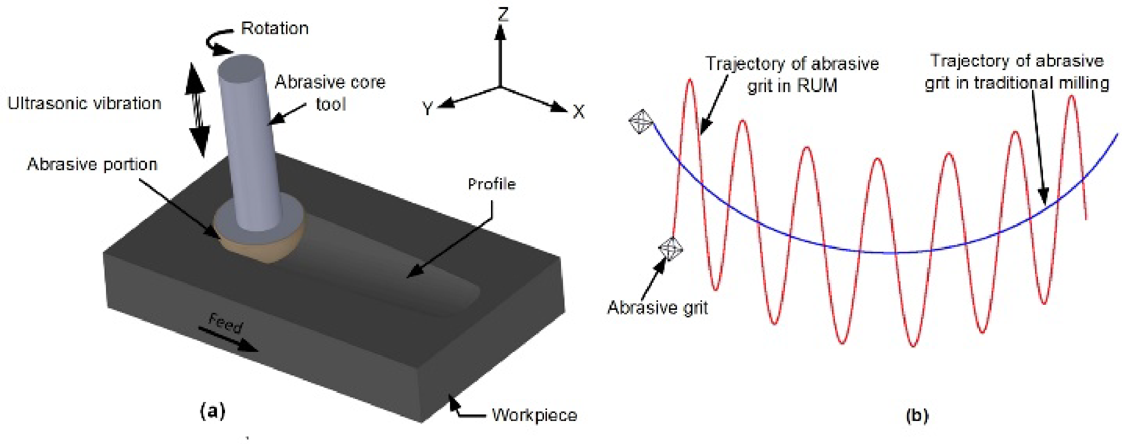


Fig. 1 Rotary ultrasonic profile milling process

Table 1 Parameters and variables applied in research work

Symbol/Abbreviation	Nomenclature
RUM	Rotary ultrasonic machining
RUPM	Rotary ultrasonic profile milling
CMC	Ceramic matrix composites
PMC	Polymer matrix composites
MMC	Metal matrix composites
MRR	Material removal rate of single abrasive grit, m^3/s
C/SiC	Carbon fiber-reinforced silicon matrix composites
β	Half angle of a diamond abrasive grit, 45°
S_a	The side length of a diamond abrasive grit, mm
w	Penetration depth of a diamond abrasive grit, mm
d	Penetration width of a diamond abrasive grit, mm
S	Spindle speed, rpm
a_p	Cutting depth, mm
f_r	Feed rate, mm/min
A	Ultrasonic vibration amplitude, 1×10^{-5} m
f	Ultrasonic vibration frequency, 20000 Hz
Z	Trajectory of a diamond abrasive grit, mm
Δt	Adequate contact time of penetration of grit in workpiece material, s
A_o	Area of the spherical abrasive core tool involved in cutting, mm
A	The radius of the spherical cap of the abrasive core tool, mm
h	Height of spherical cap of abrasive core tool, mm
R	The spherical radius of the abrasive core tool, mm
C_a	Abrasive concentration, mm
N_α	Number of active abrasive grits
C_l	Length of lateral crack, mm
C_h	Height of lateral crack, mm
d_c	Depth/ height of abrasive grit, mm
l_c	Abrasive grit length per vibration cycle, mm
l_{sec}	Chord length of segment/sector of abrasive grits (involved in cutting), mm
t	Adequate cutting time, s
K	Proportionality parameter for cutting force model
F_n	Cutting force of an abrasive grit on workpiece material, N
F_r	Radius cutting force by single grit, N
F_t	Tangential cutting force by single grit, N
F_f	Feed cutting force by single grit, N
$F_{f(m)}$	Cutting force measured from experiments, N
$F_{f(s)}$	Cutting force simulated from model without K, N
$F'_{f(s)}$	Cutting force simulated from model with K, N
V_r	Material removal volume by abrasive grit in one rotation cycle, m^3
V_r'	Material removal volume by a side face of active abrasive grit in one cycle, m^3
V_a	Actual material volume in one rotation cycle, m^3
θ	The angle between force F and cutting force F_n, (°)
v	Poisson's ratio
H_v	Vickers-hardness of the workpiece material, GPa
E	Elastic modulus, GPa
K_{IC}	Fracture toughness, $MPa\cdot m^{1/2}$
ρ	The density of workpiece material, g/cm^3
C_1, C_2, C_3	dimensionless constants

2.1 Feed-cutting force model

The feed direction cutting force prediction model is developed by considering the single abrasive grit of the core tool. The summation of all active abrasive grits is considered in cutting process. When a diamond abrasive grit penetrates the surface of workpiece, the material undergoes plastic deformation. With increased penetration depth, median and lateral cracks grow, as shown in Fig. 2. The extended lateral cracks then induce and peel off the workpiece material. The median cracks are related to degradation of strength of workpiece material, while lateral cracks are involved in material removal in the machining process of composite materials. For developing the cutting force model, maximum penetration depth has used as an intermediate parameter to establish relationships between machining and related parameters with cutting force.

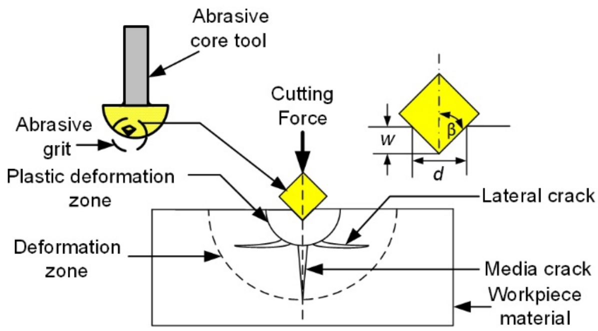


Fig. 2 Crack generation and deformation zone in material

From Fig. 2, the relationship can be established as follows:

$$w = \frac{d}{2\tan\beta} \tag{1}$$

where w is penetration depth, d is penetration width, and β is half-angle of abrasive grit ($\beta = 45°$). The volume of single diamond abrasive grit, v can be expressed as follows:

$$v = \frac{\sqrt{2}}{3} S_a^3 \tag{2}$$

where S_a is the side length of diamond abrasive grit, as shown in Fig. 3(a).

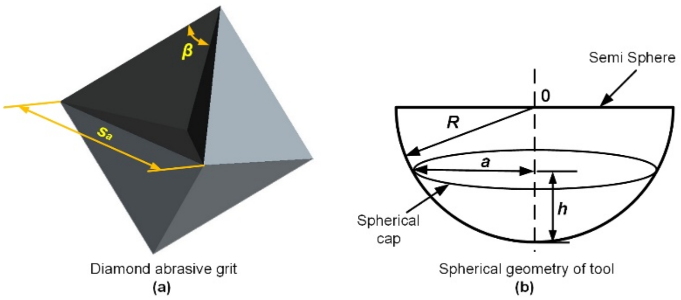


Fig. 3 Octahedron-shaped abrasive grit and related geometry

Diamond abrasive concentration in the working layer can be defined as the quantity of abrasives per unit volume. The concentration is the volume per cubic centimeter of abrasive grains containing 4.4 karats. An increase or decrease of 1.1 karats of the abrasives increases or decreases concentration by 25 %. According to this definition, the total number of active diamond abrasives/grits involved in cutting. N_α can be expressed as follows:

$$N_\alpha = \left(\frac{0.88 \times 10^{-3}}{(\sqrt{2}/3)\ S_a^3\ \rho}\frac{C_\alpha}{100}\right)^{2/3} A_0 = C_1 \frac{C_\alpha^{2/3}}{S_a^2} A_0 \tag{3}$$

where ρ is the density of diamond (3.52×10^{-3} g/mm^3), C_α is the diamond abrasive concentration, C_1 is a constant number, $C_1 = 3 \times 10^{-2}$ and A_0 is the area of abrasive tool in contact with the workpiece material (involved in cutting) at maximum cutting depth of the abrasive core tool. Since the spherical abrasive core tool is used in this study, the area of spherical cap of abrasive core tool is involved in cutting operation. The surface area of a closed spherical cap can be expressed as follows:

$$A_0 = \pi(a^2 + h^2) \tag{4}$$

where A_0 is the surface area of a closed spherical cap, a is the base radius circular sector, h is the height of the spherical cap, and R is the spherical radius of the abrasive core tool, as shown in Fig. 3(b). h shows the spherical sector's height from the base of spherical abrasive core tool. The radius of the circular sector is expressed as follows:

$$a = \sqrt{h\,(2R - h)} \tag{5}$$

From Eqs. 4 and 5, the effective surface area of the spherical core tool can be established as follows:

$$A_0 = 2\,\pi\,h\,R \tag{6}$$

Since the height of spherical cap of the tool is the height of spherical core tool involved in cutting process is the cutting depth a_p (i.e., $h \approx a_p$), the effective surface area of spherical core tool involved in machining can be expressed as follows:

$$A_0 = 2\,\pi\,a_p\,R \tag{7}$$

From Fig. 4, the relation between Z and f can be obtained as follows:

$$Z = A\sin(2\,\pi f t) \tag{8}$$

where Z represents the grain's trajectory, A and f are the magnitude and frequency, respectively, and t is the time.

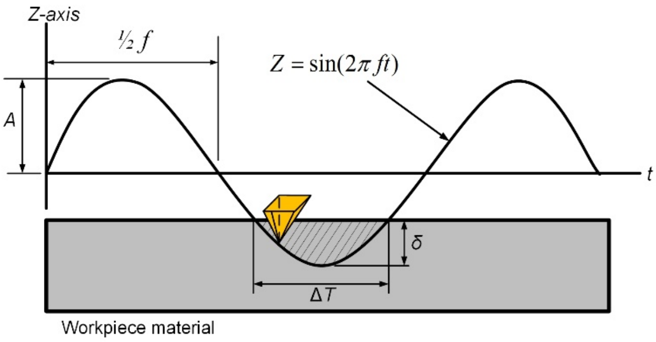


Fig. 4 Relation of adequate contact time (Δt) and maximum penetration depth (w)

During the machining, the cutting force in the feed direction is more than the axial direction; the feed direction cutting forces must be within acceptable limits. Developing a feed-cutting force model is essential for the prediction/ control of cutting forces for desired quality and reducing machining defects. The radial cutting force (F_r) and tangential force (F_t) for single diamond grit on the surface of abrasive core tool are shown in Fig. 5.

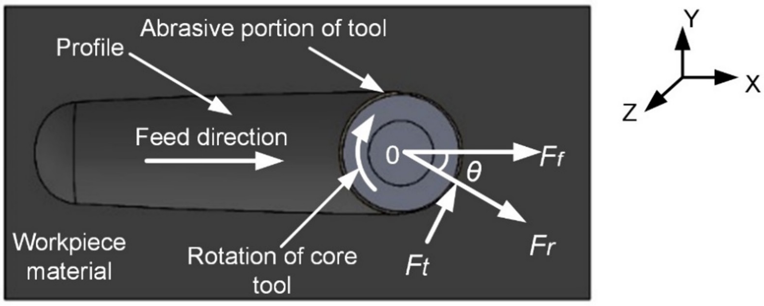


Fig. 5 Illustration of feed cutting force

The adequate cutting time in one rotation cycle is half of the cycle time, and can be expressed as follows:

$$t = \frac{1}{2}\,\frac{60}{S} = \frac{30}{S} \tag{9}$$

For cutting force in the feed direction, the active abrasive grits is expressed as follows:

$$N_\alpha = \left(\frac{0.88\ \times 10^{-3}}{(\sqrt{2}/3)\,S_\alpha^3\,\rho}\,\frac{C_\alpha}{100}\right)^{2/3} l_c\,d_c \tag{10}$$

where l_c is the length of abrasive grit that travels in one vibration cycle, and d_c is the depth/height of abrasive grit (from the bottom of the spherical abrasive core tool).

Eq. 10 can be simplified by applying the factor $\left(\frac{0.88\times10^{-3}}{(\sqrt{2}/3)\,S_a^3\,\rho}\,\frac{C_\alpha}{100}\right)^{2/3} = \frac{C_1 C_\alpha^{2/3}}{S_a^2}$
where $C_1 = \left(\frac{0.88\times10^{-3}}{(\sqrt{2}/3)\,\rho}\,\frac{1}{100}\right)$
Eq. 10 can be expressed as follows:

$$N_\alpha = \frac{C_1 C_\alpha^{2/3}}{S_a^2}\, l_c\, d_c \tag{11}$$

The factor $l_c\, d_c$ can be found by integrating (for side abrasive grits) from 0 to h for spherical abrasive core tool:

$$N_\alpha = \left(C_1 \frac{C_\alpha^{2/3}}{S_a^2}\right)\frac{1}{2}\left(\int_0^h \sqrt{(2R-h)h}\;dh\right) \tag{12}$$

$$N_\alpha = \left(C_1 \frac{C_\alpha^{2/3}}{S_a^2}\right)\frac{1}{2}\left(\frac{\pi\, R^2}{4} - \frac{1}{2}\left((R-h)\sqrt{R^2-(R-h)^2} + R^2 \arctan\left(\frac{h}{\sqrt{R^2-(R-h)^2}}\right)\right)\right) \tag{13}$$

$$N_\alpha = \frac{1}{4}\left(C_1 \frac{C_\alpha^{2/3}}{S_a^2}\right)\left(\frac{\pi\, R^2}{2} - \left((R-h)\sqrt{R^2-(R-h)^2} + R^2 \arctan\left(\frac{h}{\sqrt{R^2-(R-h)^2}}\right)\right)\right) \tag{14}$$

$$l_{sec} = \frac{\pi\, R^2}{2} - \left((R-h)\sqrt{R^2-(R-h)^2} + R^2 \arctan\left(\frac{h}{\sqrt{R^2-(R-h)^2}}\right)\right) \tag{15}$$

$$N_\alpha = \frac{1}{4} C_1 \frac{C_\alpha^{2/3}}{S_a^2}\, l_{sec} \tag{16}$$

Also, the feed direction cutting force can be expressed as follows:

$$F_f = \int F_r \sin\theta \; d\, N_\alpha \tag{17}$$

where θ is the angle between the feed cutting and radial cutting forces.

$$F_f = 2\int_0^{\frac{\pi}{2}} F_r \sin\theta \; d\, N_\alpha \tag{18}$$

By solving Eq. 18, the following relation can be obtained:

$$F_f = \frac{2\, F_r}{4} N_\alpha \tag{19}$$

where F_r is the radial cutting force by a single abrasive grit on the face of a spherical abrasive core tool, and F_n is the impact force exerted by a single abrasive grit ($F_r \approx F_n$). Eq. 19 can be expressed as follows:

$$F_f = \frac{F_n}{2} N_\alpha \tag{20}$$

The material volume removed by an abrasive grit in one rotation cycle, V_r can be found as follows:

$$V_r = 2\; C_l\; C_h\, \pi \left[\int_0^h \sqrt{(2R-h)h}\; dh\right] \tag{21}$$

By considering the number of active abrasive grits on the spherical face, the material removal volume by spherical face of active abrasive grits in one rotation cycle V_r' can be expressed as follows:

$$V_r = 2\, C_l\, C_h\, \pi \left(\frac{\pi\, R^2}{4} - \frac{1}{2}\left((R-h)\sqrt{R^2-(R-h)^2} + R^2 \arctan\left(\frac{h}{\sqrt{R^2-(R-h)^2}} \right) \right) \right) \tag{22}$$

The material removed by an abrasive grit in one rotation cycle is expressed as follows:

$$V_r = 2\, C_l\, C_h\, \pi\, \frac{1}{2}\, l_{sec} \tag{23}$$

$$V_r = \pi C_l\, C_h\, l_{sec} \tag{24}$$

The material removed by the side face of active abrasive grits in one rotation is given by:

$$V_r' = N_\alpha\, V_r \tag{25}$$

$$V_r' = N_\alpha\, \pi\, C_l\, C_h\, l_{sec} \tag{26}$$

The actual material removal in one rotation cycle is calculated by:

$$V_a = 2\, \pi\, R\, h \frac{60}{S} f_r \tag{27}$$

The relationship between V_r' and V_a is given by:

$$V_r{}' = K'\, V_a \tag{28}$$

where K' is a constant and can be found mechanistically from cutting force experiments. By putting values of $V_r{}'$ and V_a from Eq. 26 and Eq. 27, the following relation is obtained:

$$N_\alpha\, C_l\, C_h\, \pi\, l_{sec} = K' \frac{2\, \pi\, R\, h\, 60\, f_r}{S} \tag{29}$$

According to the indentation theory proposed by Marshall and Lawn [32, 33], the lateral crack length C_l and the depth C_h can be expressed as follows:

$$C_1 = C_2 \left(\frac{1}{\tan\beta}\right)^{5/12} \left(\frac{E^{3/4}}{H_v K_{IC} (1-v^2)^{1/2}}\right)^{1/2} F_n^{5/8} \tag{30}$$

$$C_h = C_2 \left(\frac{1}{\tan\beta}\right)^{1/3} \frac{E^{1/2}}{H_v} F_n^{1/2} \tag{31}$$

where E is elastic modulus, ν is the Poisson's ratio of the workpiece material, and C_2 is a dimensionless constant number, $C_2 = 0.226$ [32, 33].

Using the values of C_l and C_h in Eq. 30 and replacing Eq. 31 with Eq. 29, the following relation is obtained:

$$F_n = (K')^{8/9} \frac{(120)^{8/9}\, R^{8/9}\, H_v^{4/3}\, (1-\nu^2)^{2/9}\, K_{IC}^{4/9}\, h^{8/9}\, f_r^{8/9} (\tan\beta)^{2/9}}{S^{8/9}\, C_2^{16/9}\, E^{7/9}\, l_{\text{sec}}^{8/9}\, N_\alpha^{8/9}} \tag{32}$$

Putting the value of F_n from Eq. 32, Eq. 20 can be expressed as follows:

$$F_f = (K')^{8/9} \frac{(120)^{8/9}\, R^{8/9}\, H_v^{4/3} (1-\nu^2)^{2/9}\, K_{IC}^{4/9}\, h^{8/9}\, f_r^{8/9}\, (\tan\beta)^{2/9}\, N_\alpha^{1/9}}{2\, S^{8/9}\, C_2^{16/9}\, E^{7/9}\, l_{sec}^{8/9}} \tag{33}$$

By putting the value of N_α from Eq. 16, Eq. 33 can be expressed as follows:

$$F_f = (K')^{8/9}\ \frac{(120)^{8/9}\, R^{8/9}\, H_v^{4/3} (1-\nu^2)^{2/9}\, K_{IC}^{4/9}\, h^{8/9}\, f_r^{8/9} (\tan\beta)^{2/9}\, C_1^{1/9}\, C_\alpha^{2/27} l_{sec}}{2\, S^{8/9}\, C_2^{16/9}\, E^{7/9}\, l_{sec}^{8/9}\, S_a^{2/9}} \tag{34}$$

$$F_f = K \frac{C_3\, R^{8/9}\, H_v^{4/3} (1-\nu^2)^{2/9}\, K_{IC}^{4/9}\, h^{8/9}\, f_r^{8/9} (\tan\beta)^{2/9}\, C_\alpha^{2/27}}{S^{8/9}\, E^{7/9}\, l_{sec}^{7/9}\, S_a^{2/9}} \tag{35}$$

where C_3 is a constant and K is the proportionality parameter.

$$C_3 = \frac{(120)^{8/9}\, C_1^{\;1/9}}{2\, C_2^{\;16/9}} \qquad C_1 = 3\times10^{-2} \qquad C_2 = 0.226 \text{ [32, 33]}$$

Eq. 35 is the desired feed direction cutting force prediction model for rotary ultrasonic profile milling.

3. Experimental procedure

Experiments were conducted with different machining parameters to obtain the proportionality parameter K for validating feed direction cutting force model.

3.1 Experimental setup and conditions

The experimental rotary ultrasonic profile machining of C/SiC composite materials was conducted by using the experimental setup, as shown in Fig. 6. This setup comprises three parts, including the ultrasonic vibration system, CNC vertical machining center, and diamond abrasive spherical core tool. The ultrasonic vibration system contains an ultrasonic spindle and ultrasonic generator. The ultrasonic generator produces an ultrasonic frequency signal and provides to ultrasonic vibration spindle by producing ultrasonic vibrations with a specified amplitude. The ultrasonic vibration device containing an ultrasonic vibration spindle was fitted with a CNC vertical machining center (VMC 0850B, Shenyang, China). The cutting force was measured with a dynamometer (9257B, Kistler). The main specifications of the machine tool are given in Table 2. The mechanical properties of the workpiece material of C/SiC composites are in Table 3. The diamond abrasive spherical core tool parameters are mentioned in Table 4. The average grit size of 213 µm is calculated from supper abrasives (mesh size of 60/80 abrasive grits). The amplitude is kept on the higher side (10 µm), and ultrasonic frequency of 20000 Hz for optimum results (obtained through random experiments). The concave profile was selected for experimental machining with appropriate parameters. These are given in Table 5.

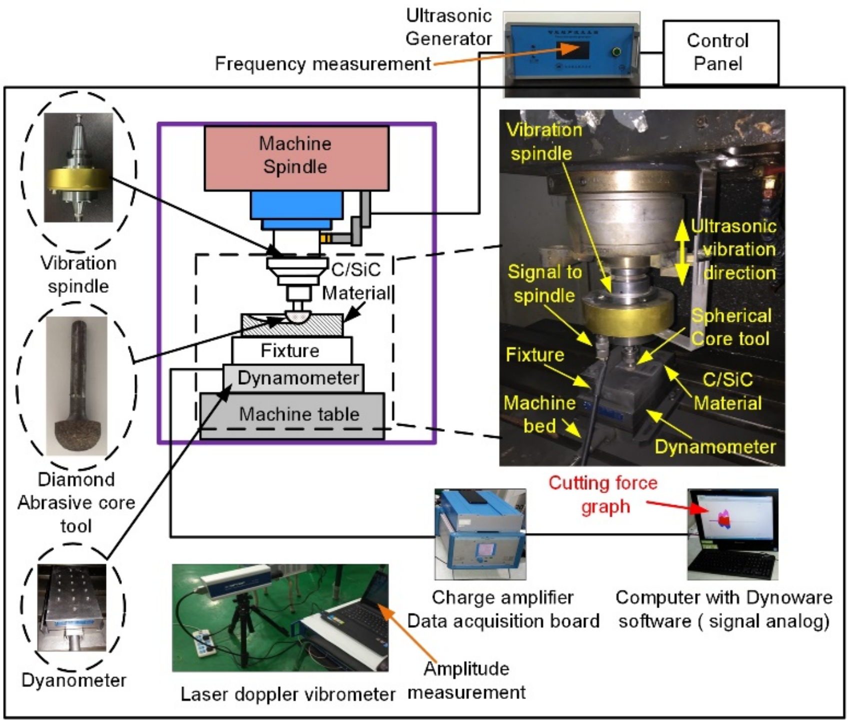


Fig. 6 Schematic along with actual setup for experiments

Table 2 Properties of the machine tool

Nomenclature	Specification
Spindle speed (with ultrasonic device)	0-6000 rpm
Ultrasonic amplitude (A)	10 µm
Ultrasonic frequency (f)	20000 Hz
Power consumption (Pw)	99 %

Table 3 Mechanical properties of C/SiC material

Nomenclature	Specification
Density (ρ)	2.0 g/cm^3
Porosity (P)	17-20 %
Tensile strength (σt)	≥ 40 MPa
Surface shear strength (σc)	≥ 10 MPa
Compression Strength(σy)	590 MPa
Elastic modulus (E)	67.7 GPa
Fracture toughness (K_{IC})	17.9 MPa·m$^{1/2}$
Vickers-hardness (H_v)	9.7GPa

Table 4 Properties of the abrasive core tool

Nomenclature	Specification
Tool type	Spherical
Abrasive	Diamond
Bond type	Metal-bond
Mesh size	60/80
Concentration (C_α)	100
Spherical radius (R)	8.25 mm

3.2 Experimental design

This study used important machining parameters, such as spindle speed, cutting depth, and feed rate based on various experimental observations for estimating effective cutting force. The experiments are designed in a single-factor experiment array with 3 factors. The level of each factor/parameter is selected by theoretical calculations, considering higher material removal rates and random experiments. The experimental design is given in Table 5.

Table 5 Experimental Design

Group	Experiments	Spindle speed S (rpm)	Feed rate f_r (mm/min)	Cutting depth a_p (mm)
1	1-7	1500, 2000, 2500, 3000, 3500, 4000,4500	100	1.0
2	7-12	3000	50,75, 100, 125, 150, 175	1.0
3	13-19	3000	60	0.7, 0.8, 0.9, 1.0, 1.1, 1.2

4. Experimental results and discussion

4.1 Measured cutting force

The experimental machining was conducted by selecting machining parameters corresponding to each group in the experimental design. The machining process is divided into three stages, i.e., enter, stable, and exit, corresponding to cutting force data (in graphical form), as shown in Fig. 7. Since profile machining is conducted in this study, the cutting force demonstrated peak values at maximum cutting depth. Therefore, the interval is considered for finding peak values of cutting force (as shown in Fig. 7). The cutting force value is the mean value of maximum values during the interval for peak value form obtained through graphical measurement with Dynoware software. The graphical cutting force data was transformed into numerical data through programming code developed in MATLAB software. The cutting force values obtained from experimental machining are shown in Table 6, corresponding to each group of parameters.

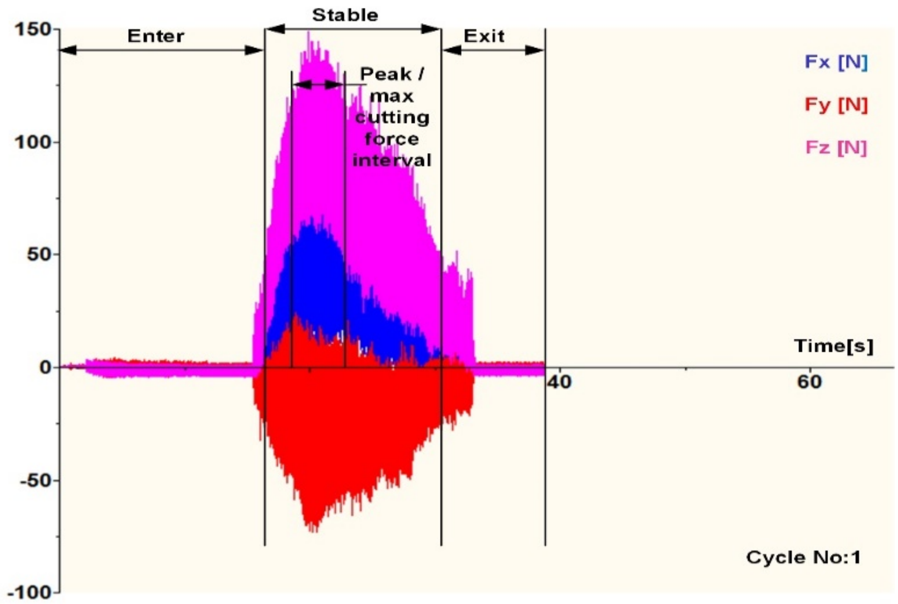


Fig. 7 Cutting force measurements (S = 2500 rpm, f_r = 100 mm/min, a_p = 1.0 mm)

4.2 The proportionality parameters

The simulated values of feed cutting force obtained through the cutting force prediction model are close to the measured cutting force when the factor gives the minimum value. The linear least square method was applied to find the value of K by partially differentiating the factor concerning K as follows:

$$\sum 2(F_{f(m)} - K\, F_{f(s)})(-F_{f(s)}) = 0 \tag{36}$$

By selecting the experimental and simulated feed cutting forces for each experiment group, the value of K was obtained as 34.842. This value gives the relationship between K and machining parameters. The simulated cutting force obtained with the feed cutting force model is given in Table 6.

Table 6 Measured and simulated feed force data

Exp. No.	S (rpm)	F_r (mm/min)	a_p (mm)	Measured feed force $F_{f(m)}$ (N)	Simulated feed force $F_{f(s)}$ without K (N)	Simulated feed force $F'_{f(s)}$ with K (N)	% Variation $\frac{F'_{f(s)} - F_{f(m)}}{F_{f(m)}}$ 100 %
1	1500	100	1.0	52.811	1.3498	53.998	+2.240
2	2000	100	1.0	50.720	1.3400	50.172	-1.080
3	2500	100	1.0	48.773	1.3324	46.423	-4.811
4	3000	100	1.0	45.615	1.3263	46.2109	+1.306
5	3500	100	1.0	42.165	1.3011	45.3329	+7.513
6	4000	100	1.0	38.508	1.2166	42.3887	+10.077
7	4500	100	1.0	43.140	1.3127	45.7370	+6.019
8	3000	50	1.0	36.426	1.1320	39.4411	+8.277
9	3000	75	1.0	45.263	1.2419	43.2702	-4.402
10	3000	100	1.0	47.263	1.3263	46.2109	-2.226
11	3000	125	1.0	48.816	1.3957	48.6289	-0.383
12	3000	150	1.0	52.763	1.4551	50.6985	-3.912
13	3000	175	1.0	59.763	1.5073	52.5173	-12.124
14	3000	100	0.7	41.127	1.3143	45.7928	+11.344
15	3000	100	0.8	42.323	1.3187	45.9461	+8.560
16	3000	100	0.9	45.754	1.3227	46.0855	+0.724
17	3000	100	1.0	46.023	1.3263	46.2109	-0.408
18	3000	100	1.1	46.599	1.3296	46.3259	-0.586
19	3000	100	1.2	50.431	1.4426	50.2630	-0.333

4.3 Analysis of measured and simulated cutting forces

The feed direction cutting forces obtained through experiments and simulated from the cutting force model are shown in Fig. 8. From the graph, simulated values of cutting forces closely match with measured cutting forces in most model parameter groups. However, higher variations are found only in three experimental groups (Exp. No. 6, 13, and 14).

The measured and simulated feed-cutting forces corroborate with experimental parameters. However, some higher variations (more than 10 %) are found for cutting forces in Exp. 6 (10.077 %), Exp. 13 (12.124 %), and Exp. 14 (11.344 %) (Table 6). The histogram error plots are shown in Fig. 9. The mean error is 1.358, with a standard deviation of 6.003. These variations predict the heterogenic and anisotropic properties of C/SiC composites. Also, considering the micro-perspective, SiC matrix is reinforced with a multilayer of carbon fibers, which causes uneven properties in feed and cutting depth directions. The proportion of SiC and carbon fibers differ in cutting area during RUPM, including the recorded variations between simulated and measured cutting forces. The generation of high temperatures causes cutting force variations.

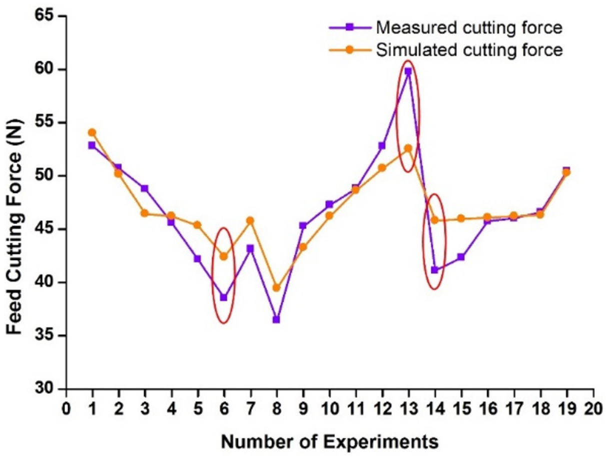


Fig. 8 Comparison of measured and simulated cutting force

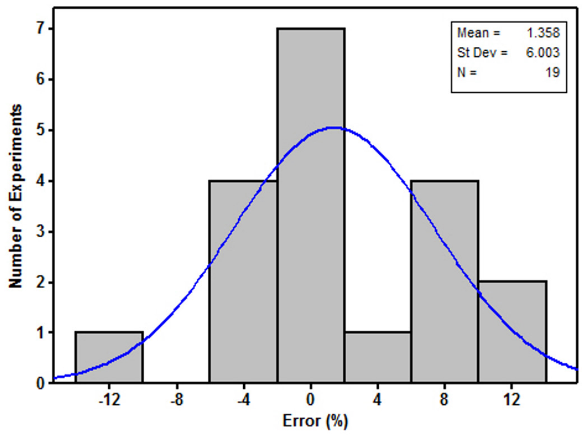


Fig. 9 Histogram plot of error vs. number of experiments

4.4 Comparison with published studies

Several studies have been reported on developing axial cutting force models for face, side, and slot milling of rotary ultrasonic machining of C/SiC composites. However, research work for the feed direction cutting force model has rarely been reported. Xiao *at al.* (2014) proposed a model for establishing cutting force during rotary ultrasonic slot milling of dental zirconia ceramics. They validated the cutting force model with max S = 6000 rpm*,* f_r = 50 mm/min, and a_p = 0.19 mm. However, the feed rate and cutting depth are significantly lower when considering material removal rates for practical applications. Li *at al.* [36] and Zhang *at al.* [37] proposed cutting force models for rotary ultrasonic face milling of C/SiC composites. The cutting force dynamic model has been proposed for rotary ultrasonic side milling of C/SiC composites.

4.5 Analysis of machining parameters

A feed rate of f_r = 175 mm/min and cutting depth of a_p = 1.2 mm is applied to develop a cutting force model. However, in published studies, such as Zhang *at al.* [29] and Xiao *at al.* [35] applied lower feed rate and cutting depth (f_r =12 mm/min, a_p = 0.08 mm, and f_r = 50 mm/min, a_p = 0.190 mm, respectively). This study used higher machining parameters to increase MRR and process efficiency for industrial applications. The feed-cutting forces found decreased with increased spindle speed. In contrast, the feed cutting forces found increased with the increase of feed rate and cutting depth. The relationship of feed cutting force with other machining parameters is shown in Fig. 10.

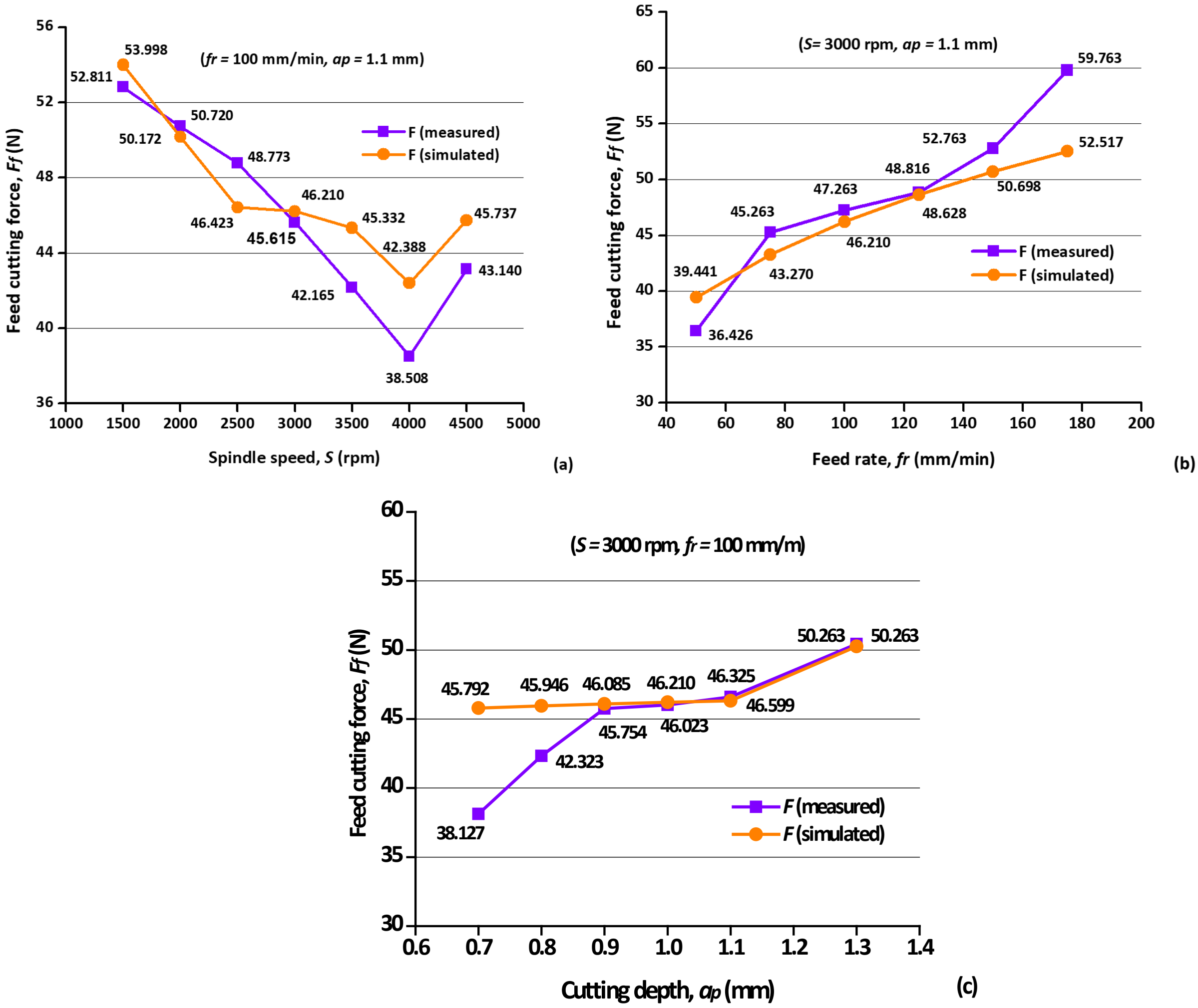


Fig. 10 Relationship of feed cutting force and machining parameters

5. Conclusion

In this study, rotary ultrasonic profile milling was conducted for machining C/SiC composites. The main contributions and conclusions of this research include:

- The cutting force prediction model is developed for RUPM of C/SiC composites and validated through experimental machining. The measured and simulated cutting forces closely match with each other. The mean error and standard deviation are 1.358 and 6.003, respectively. Variations of more than 10 % were recorded only for three groups of parameters due to material's heterogeneity and anisotropy. The developed feed cutting force prediction model is found robust and can be applied for prediction/control of cutting forces.
- The feed-cutting force prediction model for rotary ultrasonic profile milling of C/SiC composites is novel because no such study has been reported yet in published literature. This research work set new dimensions for the machining of composite materials.
- The higher machining parameters are applied ($f_r = 175$ mm/min, $a_p = 1.2$ mm, $S = 4500$ rpm) to achieve significantly higher material removal rates and practical machining conditions. The cutting force found decreased with the increase of spindle speed. However, the cutting force was increased with an increase in feed rate and cutting depth.
- The developed cutting force prediction model can be applied for predicting cutting forces in the feed direction, improving machined components' quality, and optimizing the machining process for rotary ultrasonic profile milling of C/SiC composites at the industry level.

Funding

This work was funded by the University of Jeddah, Jeddah, Saudi Arabia. The authors, therefore, acknowledge with thanks the University of Jeddah's technical and financial support.

Declaration of conflict of interest

The authors declare no conflict of interest.

References

[1] Chawla, K.K. (2003). *Ceramic matrix composites,* Second edition, Springer, New York USA, doi: 10.1007/978-1-4615-1029-1.
[2] Krenkel, W., Berndt, F. (2005). C/C–SiC composites for space applications and advanced friction systems, *Materials Science and Engineering: A,* Vol. 412, No. 1-2, 177-181, doi: 10.1016/j.msea.2005.08.204.
[3] Ferraiuolo, M., Scigliano, R., Riccio, A., Bottone, E., Rennella, M. (2019). Thermo-structural design of a ceramic matrix composite wing leading edge for a re-entry vehicle, *Composite Structures,* Vol. 207, 264-272, doi: 10.1016/j.compstruct.2018.09.024.
[4] Padture, N.P. (2016). Advanced structural ceramics in aerospace propulsion, *Nature Materials,* Vol. 15, 804-809, doi: 10.1038/nmat4687.
[5] Schmidt, S., Beyer, S., Knabe, H., Immich, H., Meistring, R., Gessler, A. (2004). Advanced ceramic matrix composite materials for current and future propulsion technology applications, *Acta Astronautica,* Vol. 55, No. 3-9, 409-420, doi: 10.1016/j.actaastro.2004.05.052.
[6] An, Q., Chen, J., Ming, W., Chen, M. (2021). Machining of SiC ceramic matrix composites: A review, *Chinese Journal of Aeronautics,* Vol. 34, No. 4, 540-567, doi: 10.1016/j.cja.2020.08.001.
[7] Teti, R. (2002). Machining of composite materials, *CIRP Annals,* Vol. 51, No. 2, 611-634, doi: 10.1016/S0007-8506(07)61703-X.
[8] Geng, D., Zhang, D., Li, Z., Liu, D. (2017). Feasibility study of ultrasonic elliptical vibration-assisted reaming of carbon fiber reinforced plastics/titanium alloy stack, *Ultrasonics,* Vol. 75, 80-90, doi: 10.1016/j.ultras.2016.11.011.
[9] Davim, J.P., Reis, P. (2005). Damage and dimensional precision on milling carbon fiber-reinforced plastics using design experiments, *Journal of Materials Processing Technology,* Vol 160, No. 2, 160-167, doi: 10.1016/j.jmatprotec.2004.06.003.
[10] Jawahir, I.S., Brinksmeier, E., , M'Saoubi, R., Aspinwall, D.K., Outeiro, J.C., Meyer, D., Umbrello, D., Jayal, A.D. (2011). Surface integrity in material removal processes: Recent advances, *CIRP Annals,* Vol. 60, No. 2, 603-626, doi: 10.1016/j.cirp.2011.05.002.
[11] Ding, K., Fu, Y., Su, H., Chen, Y., Yu, X., Ding, G. (2014), Experimental studies on drilling tool load and machining quality of C/SiC composites in rotary ultrasonic machining, *Journal of Materials Processing Technology,* Vol. 214, No. 12, 2900-2907, doi: 10.1016/j.jmatprotec.2014.06.015.
[12] Lau, W.S., Wang, M., Lee, W.B. (1990), Electric discharge machining of carbon fibre composite materials, *International Journal of Machine Tools and Manufacture,* Vol. 30, No. 2, 297-308, doi: 10.1016/0890-6955(90)90138-9.
[13] Kumar, J., Khamba, J.S. (2010). Modeling the material removal rate in ultrasonic machining of titanium using dimensional analysis, *International Journal of Advanced Manufacturing Technology,* Vol. 48, 103-119, doi: 10.1007/s00170-009-2287-1.
[14] Groover, M.P., (2010). *Fundamentals of modern manufacturing: Materials, processes, and systems,* Fourth edition, John Wiley & Sons, Hoboken, New Jersey, USA.
[15] Legge, P. (1966). Machining without abrasive slurry, *Ultrasonics,* Vol. 4, No. 3, 157-162, doi: 10.1016/0041-624X(66)90123-5.
[16] Pei, Z.J., Ferreira, P.M., Haselkorn, M. (1994). Rotary ultrasonic drilling and milling of ceramics, In: *Proceedings of the Design for Manufacturability and Manufacture of Ceramic Components Symposium, 96th Annual Meeting of the American Ceramic Society,* Indianapolis, USA, 1-12.
[17] Prabhakar, D. (1992). *Machining advanced ceramic materials using rotary ultrasonic machining process,* MS Thesis, University of Illinois Urbana-Champaign, Urbana, USA.
[18] Pei, Z.J., Ferrreira, P.M., Haselkorn, M. (1995). Plastic flow in rotary ultrasonic machining of ceramics, *Journal of Materials Processing Technology,* Vol. 48, No. 1-4, 771-777, doi: 10.1016/0924-0136(94)01720-L.
[19] Wang, J., Zhang, J., Feng, P. (2017). Effects of tool vibration on fiber fracture in rotary ultrasonic machining of C/SiC ceramic matrix composites, *Composites Part B: Engineering,* Vol. 129, 233-242, doi: 10.1016/j.compositesb.2017.07.081.
[20] Wang, J., Feng, P., Zheng, J., Zhang, J. (2016). Improving hole exit quality in rotary ultrasonic machining of ceramic matrix composites using a compound step-taper drill, *Ceramics International,* Vol. 42, No. 12, 13387-13394, doi: 10.1016/j.ceramint.2016.05.095.
[21] Pei, Z.J., Ferreira, P.M. (1999). An experimental investigation of rotary ultrasonic face milling, *International Journal of Machine Tools and Manufacture,* Vol. 39, No. 8, 1327-1344, doi: 10.1016/S0890-6955(98)00093-5.

[22] Hocheng, H., Tai, N.H., Liu., C.S. (2000). Assessment of ultrasonic drilling of C/SiC composite material, *Composites Part A: Applied Science and Manufacturing,* Vol. 31, No. 2, 133-142, doi: 10.1016/S1359-835X(99)00065-2.

[23] Li, Z.C., Jiao, Y., Deines, T.W., Pei, Z.J., Treadwell, C. (2005). Rotary ultrasonic machining of ceramic matrix composites: Feasibility study and designed experiments, *International Journal of Machine Tools and Manufacture,* Vol. 45, No. 12-13, 1402-1411, doi: 10.1016/j.ijmachtools.2005.01.034.

[24] Li, Z.C., Jiao, Y., Deines, T.W., Pei, Z.J., Treadwell, C. (2005). Rotary ultrasonic machining of ceramic matrix composites: Feasibility study and designed experiments, *International Journal of Machine Tools and Manufacture,* Vol. 45, No. 12-13, 1402-1411, doi: 10.1016/j.ijmachtools.2005.01.034.

[25] Bertsche, E., Ehmann, K., Maluhin, K. (2013). Ultrasonic slot machining of a silicon carbide matrix composite, *International Journal of Advanced Manufacturing Technology,* Vol. 66, 1119-1134, doi: 10.1007/s00170-012-4394-7.

[26] Ding, K., Fu, Y., Su, H., Cui, F., Li, Q., Lei, W., Xu, H. (2017). Study on surface/ subsurface breakage in ultrasonic assisted grinding of C/SiC composites, *International Journal of Advanced Manufacturing Technology,* Vol. 91, 3095-3105, doi: 10.1007/s00170-017-0012-z.

[27] Yuan, S., Li, Z., Zhang, C., Guskov, A. (2018). Research into the transition of material removal mechanism for C/SiC in rotary ultrasonic face machining, *International Journal of Advanced Manufacturing Technology,* Vol. 95, 1751-1761, doi: 10.1007/s00170-017-1332-8.

[28] Yuan, S., Zhang, C., Hu, J. (2014). Effects of cutting parameters on ductile material removal mode percentage in rotary ultrasonic face machining, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture,* Vol. 229, No. 9, 1547-1556, doi: 10.1177/0954405414548497.

[29] Zhang, C., Yuan, S., Amin, M., Fan, H., Liu, Q. (2015). Development of a cutting force prediction model based on brittle fracture for C/SiC in rotary ultrasonic facing milling, *International Journal of Advanced Manufacturing Technology,* Vol. 85, 573-583, doi: 10.1007/s00170-015-7894-4.

[30] Yuan, S., Fan, H., Amin, M., Guo, M. (2016). A cutting force prediction dynamic model for side milling of ceramics matrix composites C/SiC based on rotary ultrasonic machining, *International Journal of Advanced Manufacturing Technology,* Vol. 86, 37-48, doi: 10.1007/s00170-015-8099-6.

[31] Bertsche, E., Ehmann, K., Maluhin, K. (2013). An analytical model of rotary ultrasonic milling, *International Journal of Advanced Manufacturing Technology,* Vol. 65, 1705-1720, doi: 10.1007/s00170-012-4292-z.

[32] Marshall, D.B., Lawn, B.R., Evans, A.G. (1982). Elastic/plastic indentation damage in ceramics: The lateral crack system, *Journal of the American Ceramics Society*, Vol. 65, No. 11, 561-566, doi: 10.1111/j.1151-2916.1982.tb10782.x.

[33] Lawn, B.R., Evans, A.G., Marshall, D.B. (1980). Elastic/plastic indentation damage in ceramics: The median/radial crack system, *Journal of the American Ceramics Society,* Vol. 63, No. 9-10, 574-581, doi: 10.1111/j.1151-2916.1980.tb10768.

[34] Liu, D.F., Cong, W.L., Pei, Z.J., Tang, Y.J. (2012). A cutting force model for rotary ultrasonic machining of brittle materials, *International Journal of Machine Tools and Manufacture,* Vol. 52, No. 1, 77-84, doi: 10.1016/j.ijmachtools.2011.09.006.

[35] Xiao, X., Zheng, K., Liao, W. (2014). Theoretical model for cutting force in rotary ultrasonic milling of dental zirconia ceramics, *International Journal of Advanced Manufacturing Technology,* Vol. 75, 1263-1277, doi: 10.1007/s00170-014-6216-6.

[36] Li, Z., Yuan, S., Song, H., Batako, A.D.L. (2018). A cutting force model based on kinematics analysis for C/SiC in rotary ultrasonic face machining, *International Journal of Advanced Manufacturing Technology,* Vol. 97, 1223-1239, doi: 10.1007/s00170-018-1995-9.

[37] Zhang, C., Zhang, J., Feng, P. (2013). Mathematical model for cutting force in rotary ultrasonic face milling of brittle materials, *International Journal of Advanced Manufacturing Technology,* Vol. 69, 161-170, doi: 10.1007/s00170-013-5004-z.

[38] Tahir, W., Jahanzaib, M., Raza, A. (2019). Effect of process parameters on cutting speed of wire EDM process in machining HSLA steel with cryogenic treated brass wire, *Advances in Production Engineering & Management,* Vol. 14, No.2, 143-152, doi: 10.14743/apem2019.2.317.

[39] Liu, X., Wang, J., Zhu, J., Liew, P.J., Li, C., Huang, C. (2022). Ultrasonic abrasive polishing of additive manufactured parts: An experimental study on the effects of process parameters on polishing performance, *Advances in Production Engineering & Management,* Vol. 17, No. 2, 193-204, doi: 10.14743/apem2022.2.430.

APEM journal

Advances in Production Engineering & Management
Volume 18 | Number 3 | September 2023 | pp 303–316
https://doi.org/10.14743/apem2023.3.474

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

An improved deep reinforcement learning approach: A case study for optimisation of berth and yard scheduling for bulk cargo terminal

Ai, T.[a], Huang, L.[a,*], Song, R.J.[b], Huang, H.F.[c], Jiao, F.[d], Ma, W.G.[a]

[a]School of Economics and Management, Beijing Jiaotong University, Beijing, P.R. China
[b]Experimental Center of Data Science and Intelligent Decision Making, School of Management, Hangzhou Dianzi University, Hangzhou, P.R. China
[c]CRRC Information Technology CO., LTD, P.R. China
[d]Research and Development Center, Agricultural Bank of China, Beijing, P.R. China

ABSTRACT

The cornerstone of port production operations is ship handling, necessitating judicious allocation of diverse production resources to enhance the efficiency of loading and unloading operations. This paper introduces an optimisation method based on deep reinforcement learning to schedule berths and yards at a bulk cargo terminal. A Markov Decision Process model is formulated by analysing scheduling processes and unloading operations in bulk port imports business. The study presents an enhanced reinforcement learning algorithm called PS-D3QN (Prioritised Experience Replay and Softmax strategy-based Dueling Double Deep Q-Network), amalgamating the strengths of the Double DQN and Dueling DQN algorithms. The proposed solution is evaluated using actual port data and benchmarked against the other two algorithms mentioned in this paper. The numerical experiments and comparative analysis substantiate that the PS-D3QN algorithm significantly enhances the efficiency of berth and yard scheduling in bulk terminals, reduces the cost of port operation, and eliminates errors associated with manual scheduling. The algorithm presented in this paper can be tailored to address scheduling issues in the fields of production and manufacturing with suitable adjustments, including problems like the job shop scheduling problem and its extensions.

ARTICLE INFO

Keywords:
Bulk cargo terminal;
Scheduling;
Optimisation;
Markov decision process (MDP) model;
Deep reinforcement learning;
Prioritised experience replay and softmax strategy-based dueling;
Double deep Q-network (PS-D3QN)

**Corresponding author:*
lhuang@bjtu.edu.cn
(Huang L.)

Article history:
Received 22 August 2023
Revised 5 November 2023
Accepted 7 November 2023

1. Introduction

The significance of maritime logistics has been underscored by the "2022 Maritime Review" report published by the United Nations Conference on Trade and Development (UNCTAD) [1]. Maritime shipping constitutes over 80 % of global trade and is increasingly pivotal in the global economy. Ports act as crucial intermediaries, facilitating the transfer of a substantial volume of goods through loading and unloading operations. With the continuous growth of global trade and the increasing complexity of logistics, efficient port operation scheduling is vital for optimising resource utilisation, enhancing loading and unloading efficiency, and minimising operational costs.

Maritime cargo encompasses various types, including containers, dry bulk commodities (such as coal, steel, and grains), and liquids. Consequently, ports can be categorised into container terminals and bulk cargo terminals. Unlike container terminals, bulk cargo terminals involve a more intricate range of goods, each necessitating distinct loading and unloading processes. Currently, most bulk cargo terminals rely on manual expertise for scheduling, which may compromise the efficiency and rationality of scheduling plans. Furthermore, berth and yard scheduling are often treated separately in practical scheduling processes, lacking a unified approach. Thus, an intelligent berth and yard scheduling method in bulk cargo terminals is imperative to enhance logistical efficiency, on-time delivery rates, and port operational cost reduction, thereby contributing to sustainable economic development.

While a substantial body of literature exists on port resource scheduling and operational optimisation, research on optimisation for bulk cargo terminals is relatively limited compared to container terminals. The primary distinctions between container and bulk cargo terminals lie in the layout of port resources, loading and unloading procedures, handling machinery, and cargo types [2]. Given the more intricate variety of goods involved in bulk cargo terminals, it is imperative to delve into the optimisation of their loading and unloading procedures. Within bulk cargo terminals, berths and yards are the most critical and scarce resources during operational processes, making berth and yard allocation the focal points.

Berth Allocation Problem (BAP), as highlighted by Bierwirth *et al*. [3, 4] in 2010 and 2015, has been a subject of study, with subsequent research largely building upon their classification scheme. This classification is equally applicable to bulk cargo terminals. BAP can be categorised into four classes based on space, time, processing time, and performance metrics [4]. Spatially, terminals may exhibit discrete, continuous, or hybrid layouts. Temporally, BAP can be classified as static arrival, dynamic arrival, periodic arrival, and stochastic arrival. Considering the influence of ship loading and unloading times, BAP can be classified as "fixed" or "pos", depending on the position. Performance metrics categorise BAP models based on optimisation goals, with most research in BAP prioritising minimising total ship dwell time.

Berth optimisation in bulk cargo terminals aims to improve berth operational efficiency and optimise berth allocation. In a study by Đelović *et al*. [5], the authors analysed the berth productivity of multifunctional bulk terminals using mathematical and statistical methods. Their goal was to systematically identify the factors influencing berth productivity and categorise 14 groups of influential factors. The study confirmed a substantial disparity between gross and net berth productivity, primarily attributable to the substantial portion of non-operational time during vessel dwell periods.

For the BAP in bulk cargo terminals, Barros *et al*. [6] addressed the problem in tidal bulk cargo ports, formulating an integer linear programming model for discrete terminal layout and dynamic arrival scenarios while accounting for inventory constraints. Umang *et al*. [7] explored mixed terminal layouts and dynamic arrivals, proposing exact methods and a heuristic approach for minimising total service time, considering various cargo types aboard ships. Ribeiro *et al*. [8] tackled the discrete BAP in a dynamic arrival scenario for ore terminals, aiming to minimise delay and scheduling costs by employing mixed-integer linear programming and adaptive large neighbourhood search. Ernst *et al.* [9] researched continuous BAP with dynamic ship arrivals under tidal constraints, proposing two mixed-integer linear formulations and testing them on different instances. Cheimanoff *et al*. [10] studied multiple continuous terminals with dynamic arrivals, considering tidal restrictions and terminal-specific limitations for each ship, using mixed-integer linear models and iterative local search for small- and large-scale instances.

Yard management is a cornerstone of port terminal operations, as intelligent management of storage and transportation within yards can optimise space utilisation and decrease ship loading and unloading times, thereby enhancing port operational efficiency. Research focusing on yard allocation optimisation is limited compared to BAP studies. Tang *et al*. [11] examined joint storage space allocation and ship scheduling, establishing a mixed-integer programming model solved via the Benders decomposition algorithm. In their work, yard storage areas were divided into slots, each dedicated to a single product, with the possibility of extending product stacks across multiple slots. Rocha de Paula *et al*. [12] devised a genetic algorithm to maximise coal terminal throughput by arranging coal arrivals, determining stack and recovery cycles, and scheduling ship arrival and departure times.

In bulk cargo terminals, BAP is often coupled with yard allocation problems. Robenek *et al*. [13] extended the work of Umang *et al*. [7], expanding BAP to allocate yard positions to incoming ships based on specific cargo types, aiming to minimise ship service time. Unsal *et al.* [14] integrated berth allocation, stacker scheduling, and yard allocation in the context of exporting coal terminals. This problem entailed operational challenges and constraints concerning tidal windows, multiple stocking pads, non-crossing stackers, ships and berth size, addressed through integer programming formulations.

For integrated scheduling in bulk cargo terminals, some studies have combined berth and yard allocation, albeit primarily focusing on ship operation time as an optimisation goal and overlooking transport costs. Others have considered coal and ore terminals, both specialised cases, and may not be easily extended to highly diversified bulk cargo terminals. Therefore, this study focuses on a comprehensive bulk cargo terminal with diverse goods and aims to optimise berth and yard allocation in a dynamic discrete environment, minimising ship stay time and total transport costs.

Presently, research on bulk cargo port scheduling employs mathematical programming methods and intelligent algorithms, including heuristics, simulation, and genetic algorithms. These conventional mathematical and intelligent optimisation algorithms can yield favourable results when aptly modelled for certain issues. However, the berth and yard scheduling issues addressed in this study are characterised by dynamic and uncertain environments with a large scope. Traditional mathematical programming methods, heuristic algorithms, and similar optimisation techniques may lack the flexibility to address real-time changes in complex production scheduling scenarios. In contrast, recent advancements in artificial intelligence methods, such as deep learning and reinforcement learning, offer promising solutions to such challenges. For instance, Tian *et al*. [15] proposed a data prediction model for the dynamic job-shop scheduling problem (DJSP) using the Long Short-Term Memory Network (LSTM). They improved the model by integrating Dropout technology and other techniques, subsequently assessing its performance. Moreover, they devised a scheduling model with objective functions encompassing maximum makespan, total device load and key device load. Ultimately, an enhanced Multi-Objective Genetic Algorithm (MOGA) was formulated to tackle this challenge.

The scheduling problem under study falls under the category of sequential decision-making in a finite state space. The environment's state at the next time step is solely influenced by the current environment state and the actions taken by port resources. It follows the Markov property and can be formulated as a Markov Decision Process (MDP). Reinforcement Learning [16] is an artificial intelligence technique designed to address MDPs, making it well-suited for solving the scheduling problem presented in this study. Reinforcement learning algorithms have matured over recent years and span multiple branches. Depending on action selection methods, reinforcement learning can be classified into value-based methods and policy gradient-based methods [16]. Among the most common value-based algorithms are the Deep Q-Network (DQN) [17] and its variants. The Double DQN algorithm proposed by Van Hasselt *et al*. [18] and the Dueling DQN algorithm by Wang *et al*. [19] optimise target network Q-value computation and neural network architecture, respectively. Additionally, the DDPG algorithm is a popular policy gradient-based method [20]. Given that this scheduling problem involves discrete action spaces and unknown state transition probabilities, requiring the agent to continually interact with the environment for learning, value-based model-free DQN algorithms and their variants are better suited for solving this problem.

Deep learning and reinforcement learning have found applications across various fields [21-23]. However, its application to port resource scheduling remains relatively limited. Li *et al*. [24] developed a MILP mathematical model to minimise total ship stay time and employed a genetic algorithm as a fundamental optimisation method. They introduced a Q-learning approach with dynamic parameter selection for crossbreeding and mutation, along with a simulated annealing operation to address ship scheduling. Dai *et al*. [25] examined BAP and QCSP for container terminals. They created a Markov Decision Process model accounting for terminal loading capacity, cargo types, and switch setup time. Their research involved greedy insertion algorithms and DDQN reinforcement learning algorithms for offline and online scenarios. Li *et al*. [26] proposed an improved Double DQN algorithm for scheduling bulk cargo loading at a coal terminal. Their approach enhanced the ε-greedy policy and introduced a random policy for illegal actions, increasing algorithm convergence.

This study presents a deep reinforcement learning approach called the Prioritised Experience Replay and Softmax strategy-based Dueling Double Deep Q-Network (PS-D3QN). The effectiveness of this method is validated through a case study on import operations at a port in southern China. By considering the scheduling processes and unloading operations, the dynamic discrete environment of bulk terminal berth and yard scheduling is modelled as a Markov Decision Process (MDP). The PS-D3QN algorithm is subsequently employed to solve this model using actual port conditions and collected data.

The main contributions of this study are summarised as follows:

- By analysing import business scheduling processes and ship unloading operations at the port, along with incorporating real port conditions and related data, this study formulated the problem of berth and yard scheduling in bulk cargo terminals as a Markov Decision Process (MDP) model. The model's state space and action space were designed, aiming to minimise total ship stay time at the port and total transport costs. A linearly weighted reward function was devised, and legal action validity was defined, providing the basis for the subsequent introduction of deep reinforcement learning algorithms.
- This study introduced a berth and yard real-time scheduling method (PS-D3QN) based on an improved DQN algorithm. This method combined the advantages of the Double DQN and Dueling DQN algorithms, optimising the algorithm by introducing a well-designed Prioritised Experience Replay (PER) mechanism and a softmax action selection strategy. This optimisation enhanced the algorithm's convergence and stability.
- The proposed PS-D3QN algorithm was validated using actual port data from a bulk cargo terminal case study. Comparative analyses were conducted with the Double DQN and Dueling DQN algorithms, as well as real-world scheduling plans. The experimental results demonstrated the effectiveness and reliability of the proposed algorithm in addressing the bulk cargo berth and yard scheduling problem.

The rest of this paper is organised as follows. Section 2 introduces the model construction and optimisation algorithm design. Section 3 presents numerical experiments and discussions using actual port data. Finally, Section 4 summarises the study's achievements and contributions and suggests avenues for further improvement.

2. Models and algorithms

2.1 Problem statement and MDP modelling

This study focuses on the ship unloading operations within the import business of a bulk cargo terminal located in southern China. The research commences by investigating the specific operational environment and business procedures of the port. A detailed analysis of the port's actual scheduling process follows. The simplified scheduling process is illustrated in Fig. 1.

Upon receiving ship forecast information, the planning department formulates day and night plans considering current berth utilisation. Subsequently, the warehouse department develops yard operation plans based on these day and night plans and the cargo transportation mode. After the ship's arrival, the scheduling department arranges berthing, while the warehouse department coordinates ship unloading in accordance with yard operation plans and the day-night scheme. The scheduling department concludes the ship's operations by orchestrating its departure from the port.

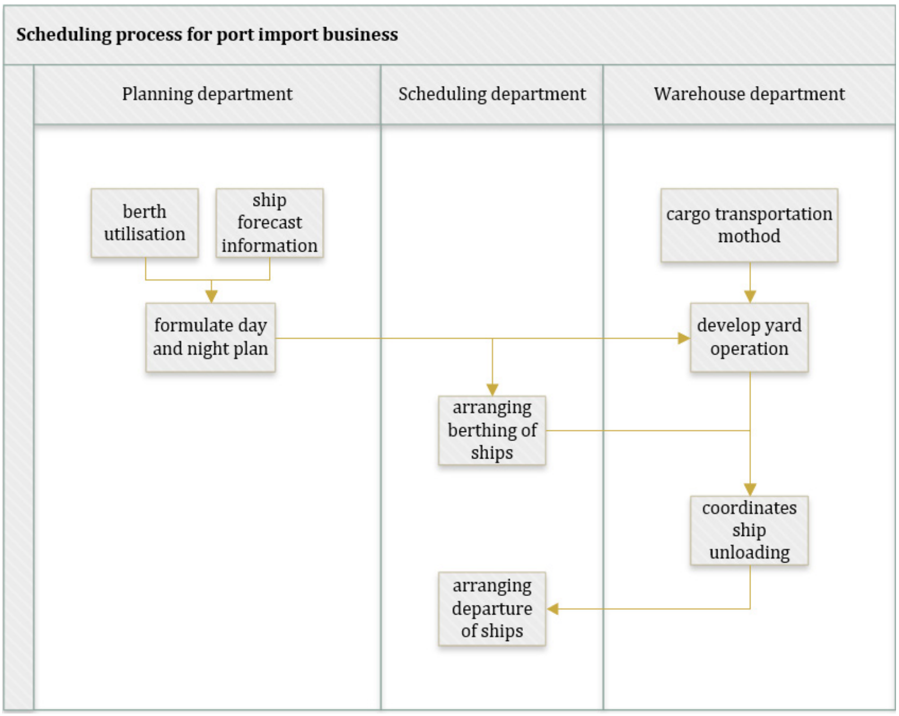


Fig. 1 Scheduling process for port import business

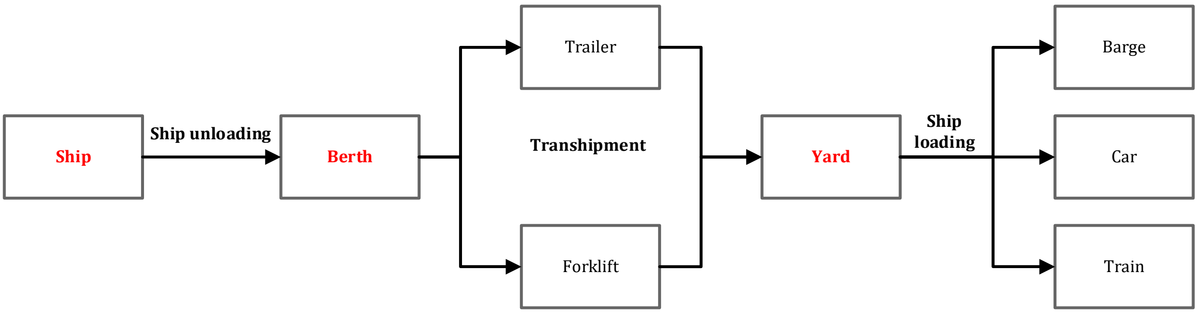


Fig. 2 Production operation process of port import business

The production operations of port import businesses encompass two main facets: ship unloading and transportation operations. Fig. 2 provides an overview of these processes. Ship unloading involves berthing, cargo unloading, transport, and storage. Conversely, transportation operations encompass the relocation of cargo within the yard and its transportation outside the port via railways, highways, coastlines, and other channels. Throughout the port's operational procedures, berths and yards, as precious resources, serve as pivotal nodes bridging ship unloading and transportation operations.

Analysing the port import business's scheduling process and its production operations highlights the current practice of separately planning berth and yard scheduling within bulk cargo terminals. This approach neglects the interdependencies and mutual constraints between yard and berth resources. Furthermore, the scheduling process relies heavily on manual experience, leading to subjective influences. Communication and coordination among departments demand considerable time, ultimately resulting in suboptimal scheduling efficiency and a lack of scientifically driven optimisation goals and decision-making criteria. Consequently, the challenge of port berth and yard scheduling necessitates a holistic resource allocation optimisation supported by intelligent methodologies to enhance scheduling efficiency, reduce manual scheduling costs, and elevate port production and operation efficiency.

This paper's scheduling problem can be defined as follows: Given the planned arrival time and essential information about ships, berths, and yards, the objective is to address the unloading operations in the port's import business. Specifically, the study encompasses all ships anticipated to arrive and depart within a fixed planning period. The focus lies on ship docking, loading/unloading operation timing, and the unloading location. The optimisation aims to minimise the total dwell time for all arriving ships and the aggregate cargo transportation costs, resulting in berth and yard allocation plans for each ship.

This study establishes an MDP model by integrating the operational processes, layout environment, and actual data of ships, berths, and yards in the port to address this scheduling problem. The pertinent design elements of the model are detailed as follows:

(1) State space
The state space encompasses the operational and usage states of ships, berths, and yards. This information is summarised in Table 1.

Table 1 State space of berth and yard scheduling in bulk terminals

State space	Data structure	Dimension	Category	Description
C_s	Arrays	10	Int	Type of cargo loaded on the ship (0: empty cargo ship, 1: steel, 2: coal, 3: grain, 4: ore)
W_s	Arrays	10	Int	Weight of cargo loaded on the ship
B_s	Arrays	10	Int	Ship docking position
S_s	Arrays	10	Int	Ship operational status (0: waiting, 1: unloading, 2: completed)
Y_s	Arrays	10	Int	Location of storage of cargo loaded on the ship
B	Arrays	6	Boolean	Whether the berth is occupied
Y	Arrays	20	Boolean	Whether the yard is occupied
C_y	Arrays	20	Int	Type of cargo stored in yard space
W_s	Arrays	20	Int	Weight of cargo stored in yard space

(2) Action Space
To tackle the scheduling problem, the action space incorporates ship berthing and cargo storage yard allocations. Each action comprises two components: berth and yard. The former denotes the berth number representing the ship's docking location, while the latter indicates the yard number for cargo storage. The action space encompasses 10 ships, 6 berths, and up to 20 stacks, amounting to a total of 1200 possible actions. However, due to varying ship arrival times and the presence of docking and storage constraints, certain actions will be infeasible. Hence, a filtering process is necessary when designing the action space.

(3) Reward function
This paper belongs to the Multi-Objective Reinforcement Learning (MORL) problem, aiming to minimise the total dwell time of all ships in port and the aggregate cargo transportation cost. Each objective corresponds to a distinct reward function. Consequently, the overall reward consists of a collection of individual objective vectors. When objectives are directly correlated (e.g., minimised time or cost), MORL can be transformed into a single-objective RL problem through linear weighting. In reality, objectives frequently feature conflicts or constraints, requiring selective optimisation or trade-offs between conflicting objectives [27]. Both objectives in this study pertain to the minimisation of total time or cost. Thus, a linear weighted approach is adopted to formulate the reward function.

The designed reward function is expressed by Eq. 1:

$$R = -k(T + S) - l(D \times W) + C \tag{1}$$

Here, k and l denote the weights of the two objectives. After empirical investigation, k is set to 0.7 and l to 0.3. T represents the ship's operation time, calculated by dividing the weight of loaded cargo by the average operation speed of the berth corresponding to the cargo type. S signifies the ship's waiting time, determined by subtracting the arrival time from the commencement of operations. D captures the total cargo distance from berth to yard. W stands for cargo weight, and C is an adjustment value. Positive rewards are bestowed on reasonable actions, whereas penalties are applied through negative reward functions for suboptimal choices. To encourage intelligent agents to select legitimate actions, penalties for illegitimate actions slightly exceed positive rewards, thus fostering effective learning.

By crafting a well-designed reward function, intelligent agents are incentivised to choose appropriate actions while avoiding improper selections, leading to enhanced learning outcomes.

2.2 Berth and yard scheduling approach based on PS-D3QN

Reinforcement learning models for berth and yard scheduling in bulk ports entail managing substantial state variables and action decisions. Moreover, their state and action spaces exhibit considerable complexity, necessitating the employment of the DQN algorithm for approximating high-dimensional states. Derived from the DQN algorithm, Double DQN and Dueling DQN are advanced techniques that address its limitations. Double DQN overcomes overestimation issues by estimating target network Q-values using the action selected based on current evaluation network Q-values. On the other hand, Dueling DQN enhances stability by decoupling action-value functions through modifying neural network structure and achieving more accurate Q-value estimation.

This paper introduces a real-time scheduling approach, termed PS-D3QN, for berths and yards, based on an enhanced DQN algorithm. PS-D3QN integrates the Q-value estimation methodology of Double DQN and the concept of action-value function separation from Dueling DQN, synergising their strengths to enhance algorithm performance. Additionally, algorithm performance is further improved through the refinement of Prioritised Experience Replay (PER) and Softmax strategies.

In the PS-D3QN framework proposed in this study, the action value function is decomposed into a combination of state value V and action advantage function A, enabling a more valuable assessment of actions. The state value reflects the current state, while the action advantage function measures the disparity between current action performance and average performance. Actions that outperform the average yield a positive advantage function, while others yield a negative advantage function. Given a fixed Q, countless combinations of V and A can generate Q. Consequently, restrictions are imposed on A; typically, the average of the advantage function A for the same state is constrained to 0. Thus, the action value function is calculated as shown in Eq. 2:

$$Q(s_t, a_t) = V(s_t) + \left(A(s_t, a_t) - \frac{1}{|A|} \sum\nolimits_{a_t} A(s_t, a_t) \right) \tag{2}$$

In PS-D3QN, the maximum action value from the evaluation network is used to calculate Q-values in the target network. The target network's value function is derived according to Eq. 3, where θ_t and ${\theta_t}^{-}$ denote the parameters of the evaluation and target networks, respectively.

$$ {Y_t}^{PS-D3QN} = R_{t+1} + \gamma Q' \left(S_{t+1}, \underset{a}{argmax}\, Q(S_{t+1}, a; \theta_t)\, ; {\theta_t}' \right) \tag{3}$$

The primary objective of PS-D3QN is to train parameters that minimise the loss function, formulated in Eq. 4.

$$L^{PS-D3QN}(\theta_t) = E\left(\left({Y_t}^{PS-D3QN} - Q_t(S_t, a; \theta_t) \right)^2 \right) \tag{4}$$

Following loss function computation, PS-D3QN employs stochastic gradient descent to update training parameters, transferring them to the target network parameters as illustrated in Eq. 5.

$$\theta_{t+1} = \theta_t + \alpha E\left({Y_t}^{PS-D3QN} - Q_t(S_t, a; \theta_t) \frac{\partial Q_t(S_t, a; \theta_t)}{\partial \theta_t} \right) \tag{5}$$

The neural network architecture of the PS-D3QN algorithm, presented in Fig. 3, is constructed based on the MDP model established in Section 2.1.

PS-D3QN employs two networks: the evaluation network and the target network. Their structures are identical, featuring an input layer, two fully connected hidden layers, and an output layer. The ReLU function serves as the neuron activation function. Input consists of the state set, with output comprising the action set. The second fully connected layer separately outputs state values and action advantage functions, combining to present individual actions and their respective Q-values.

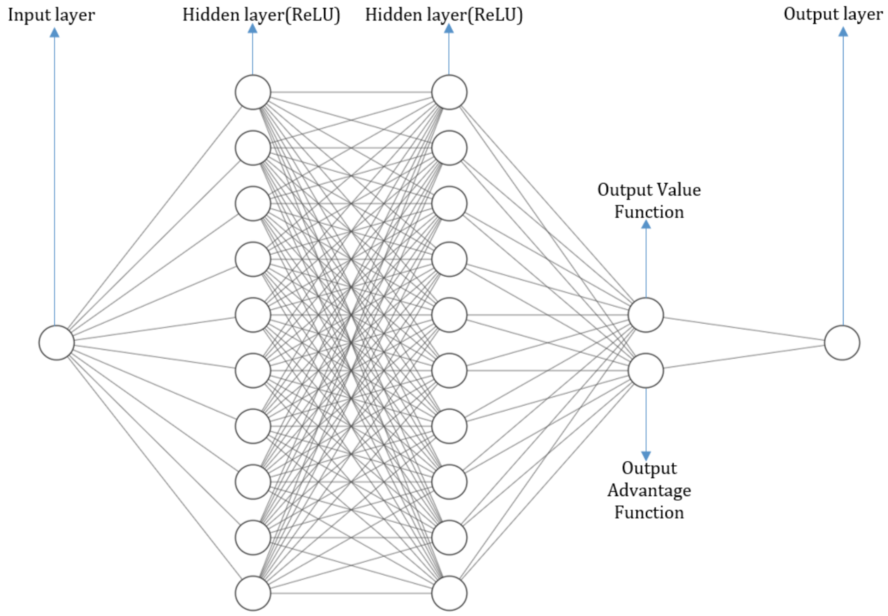


Fig. 3 Neural network structure of PS-D3QN

The PS-D3QN algorithm proposed in this study optimises its performance using the Prioritised Experience Replay (PER) mechanism, assigning priority to each experience based on Time Difference Error (TD-error). By favouring experiences with higher priority, the algorithm is inclined to select them for training. To avoid an excessive focus on high-priority experiences, this study introduces priority sampling weights for experience extraction, enhancing experience utilisation and promoting effective training and convergence. The experience priority in the algorithm, shown in Eq. 6, is influenced by TD-error and the number of completed tasks. This approach facilitates selecting valuable data for training, accelerating learning, and improving performance.

$$P = |Q_t - Q_c| + \frac{N_t}{N_t + 1/\sigma} \tag{6}$$

In Eq. 6, P represents the experience priority, Q_t stands for the target Q value, Q_c corresponds to the current Q value, N_t denotes the current number of completed tasks, σ signifies the weight, which progressively increases with the number of iterations, eventually reaching a final value of 0.01. Within this paper, experience priority is influenced by TD-error and the number of completed tasks. Consequently, the algorithm tends to favour more valuable data during training, enhancing the reuse of pivotal experiences. This approach accelerates the learning process and enhances the algorithm's overall performance.

The PS-D3QN method proposed here replaces the ε-greedy strategy with the Softmax strategy from the DQN algorithm. Softmax is a common technique for balancing exploration and exploitation in reinforcement learning, selecting actions based on a probability distribution derived from each action's estimated average reward. This strategy encourages frequent selection of actions with higher average rewards while still exploring other actions through non-zero probabilities. Softmax, governed by the Boltzmann distribution, allocates probabilities to action selection based on the estimated average reward. As depicted in Eq. 7:

$$P(k) = \frac{e^{\frac{Q(k)}{\tau}}}{\sum_{i=1}^{K} e^{\frac{Q(i)}{\tau}}} \tag{7}$$

In Eq. 7, $P(k)$ denotes the probability of selecting action k, $Q(k)$ represents the estimated average reward value for action k based on historical data, and $Q(i)$ records the average reward value after the current action's completion. τ, referred to as the "temperature", influences the trade-off between exploration and exploitation. Lower τ values emphasise exploitation, while higher values encourage exploration. In this study, τ is set using hyperbolic annealing. Initially, a higher temperature is employed to promote exploration, gradually reducing as experience accumulates to encourage utilisation of well-performing actions. The temperature update process is governed by Eq. 8, where τ_0 is the initial temperature and τ_k controls the annealing rate.

$$\tau(i) = \frac{\tau_0}{1 + \tau_k i} \tag{8}$$

The overall flow of the PS-D3QN algorithm is presented in Fig. 4.

During PS-D3QN algorithm training, the environment is initialised with yard storage information, berth occupancy data, ship states, and cargo details. Based on the current state, the Softmax strategy is used to select scheduling actions for berths and yards. Throughout the training, the algorithm sequentially stores experiences derived from interacting with the environment in the experience replay pool. Once sufficient experiences are collected, the PS-D3QN algorithm conducts random and priority sampling from the experience replay pool based on priority sampling weights.

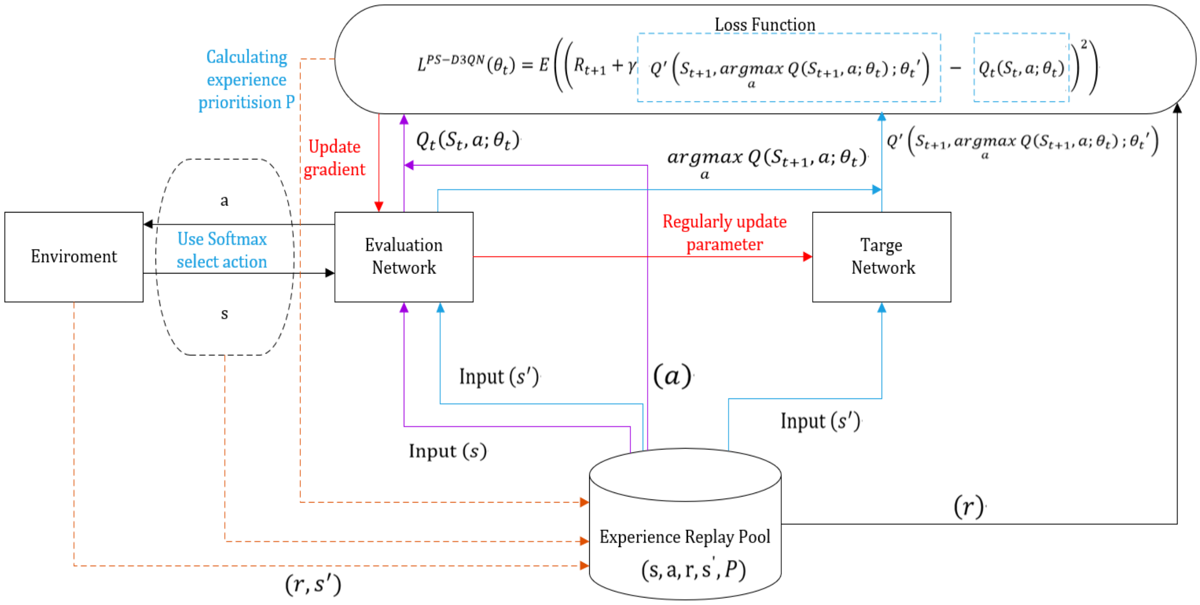


Fig. 4 PS-D3QN algorithm flowchart

3. Experimental results and discussion: A case study for bulk cargo terminal

The PS-D3QN algorithm proposed in this paper is deployed on a cloud server. The relevant environment configuration includes Windows Server 2022, Python 3.8, and PyTorch 1.12. The processor is a 16-core Intel Xeon (Ice Lake) Platinum 8369B, the GPU is NVIDIA Tesla A100 (80GB video memory), and the memory is 125GB.

This paper presents a case study on ship unloading operations in the import business of a port in southern China. In this section, we validate the proposed PS-D3QN algorithm using dynamic ship arrival data for a single planning period. The selected ship arrival data comprehensively covers most cargo types, ensuring good representativeness and generalisation. Based on this data, the PS-D3QN algorithm is trained and its results are compared and analysed against the Double DQN algorithm, Dueling DQN algorithm, and the actual scheduling scheme. The algorithm's relevant parameters are presented in Table 2.

The arriving ship data, berth data, and yard data used in this study are provided in Tables 3, 4, and 5. After conducting numerical tests, the simulated reward function curve and simulated loss function curve of the PS-D3QN algorithm, Double DQN algorithm, and Dueling DQN algorithm are illustrated in Fig. 5 and Fig. 6.

Table 2 Parameter settings

Parameter	Description	value
α	Learning Rate	0.001
γ	Discount Factor	0.99
Replay buffer size	Experience replay pool capacity	10000
Batch size	Batch size of samples per training	32
Tau	Rate at which the target network copies weights from the evaluation network	0.001
episode	Maximum number of steps per training round	500
Alpha_Prioritise	Weights for prioritised sampling	0.6
τ_0	Softmax initial temperature	200
τ_k	Softmax annealing speed	0.01

Table 3 Arriving ship data

Serial number	Ship name	Length	Depth	Cargo type	Cargo name	Cargo weight	Estimated arrival time(data preprocessing)
1	Ship 1	112.21	9	grain	cassava	11531	0d 13h
2	Ship 2	158.8	10	coal	coal	24376	0d 7h
...							
9	Ship 9	166.31	9	steel	steel	13015	0d 16h
10	Ship 10	149.18	10	grain	soya	21825	1d 23h

Table 4 Berth data

Serial number	Berth	Length	Depth	Operational speed of ore (tonnes per hour)	Operational speed of steel (tonnes per hour)	Operational speed of coal (tonnes per hour)	Operational speed of grain (tonnes per hour)
1	Berth 1	181	9	300	450	0	0
2	Berth 2	192	9	350	600	0	0
...							
5	Berth 5	201	10.5	240	0	110	180
6	Berth 6	202	10.5	360	0	500	210

Table 5 Yard data

Serial number	Yard	Cargo type	Yard type	Yard capacity	Horizontal relative position
1	Position 1 in district 1	Coal	Outside	25000	2
2	Position 2 in district 1	Coal	Outside	25000	2
3	Position 3 in district 1	Steel	Outside	20000	2
...					
19	Position 1 in district 7	Grain	Warehouse	20000	8
20	Position 2 in district 7	Grain	Warehouse	20000	8

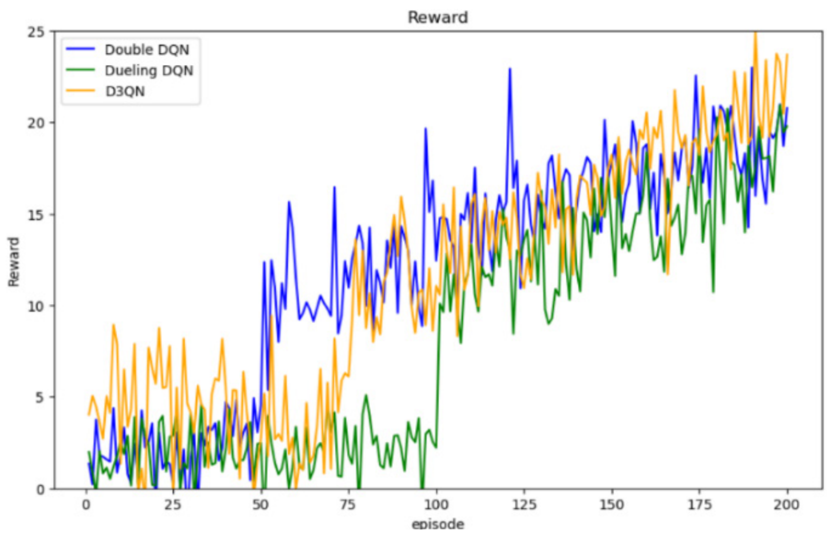


Fig. 5 Simulation reward function curve

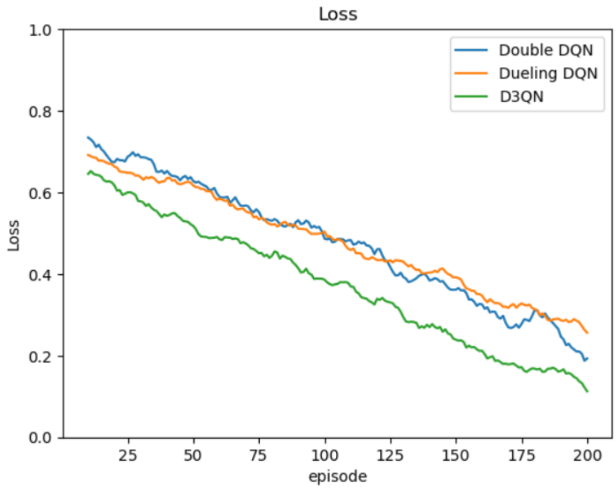


Fig. 6 Simulation reward loss curve

As shown in Fig. 5, the Double DQN algorithm quickly identifies a relatively high reward value in the initial iterations due to mitigating the overestimation problem. However, as training progresses, it might get trapped in a local optimum. Dueling DQN exhibits slower convergence because of its complex network structure, yet its decoupled action-value function ensures more accurate final computation results. The PS-D3QN algorithm proposed in this paper maintains faster convergence, stability, and superior scheduling outcomes by combining the strengths of both algorithms.

Fig. 6 portrays the decreasing trend in the simulated loss function curves for the three algorithms. The PS-D3QN algorithm's curve exhibits a smoother decrease compared to the other two algorithms. From the start to the end of iterations, its loss function gradually decreases, converging towards 0 with a relatively swift convergence rate.

To summarise, Double DQN's quick convergence is attributed to alleviating the overestimation problem. However, it can get trapped in a local optimum, leading to convergence on a locally optimal strategy. Dueling DQN's slower learning process due to increased network complexity, is associated with greater fluctuations and inadequate stability despite eventual convergence. The PS-D3QN algorithm, as proposed, excels in convergence speed and stability. It efficiently discovers maximum reward values, leveraging the combination of the other two algorithms to address overestimation issues while enhancing Q-value estimation through action value function splitting, thereby streamlining the training process. Incorporating the Prioritised Experience Replay mechanism and the Softmax action selection strategy optimises computational efficiency and stability. These enhancements facilitate faster convergence while overcoming local optima issues.

Following numerical experiments, the final scheduling results of the three algorithms and the outcomes of the actual scheduling scheme are presented in Table 6.

The PS-D3QN algorithm, Double DQN algorithm, and Dueling DQN algorithm each improve scheduling scheme efficiency by 12.85 %, 9.18 %, and 8.93 %, respectively, compared to the actual scheduling scheme. The scheduling scheme derived from the PS-D3QN algorithm effectively reduces ships' port dwell time, thereby laying the groundwork for subsequent ship arrivals. Furthermore, the scheme contributes to significant reductions in total cargo transportation costs, enhancing ship loading and unloading efficiency and subsequently reducing overall port operating costs.

Table 6 Scheduling results

Scheduling scheme	Total ship dwell time (hours)	Total costs of cargo transport (ten thousand tonnes multiplied by metres)
PS-D3QN	437	27.6
Double DQN	449	30
Dueling DQN	441	31.9
Actual scheduling scheme	464	39

The complex nature of bulk berth and yard scheduling, characterised by extensive state and action spaces, is effectively addressed by the PS-D3QN algorithm proposed in this study. The algorithm leverages deep neural networks for nonlinear modelling and approximation. By integrating the strengths of the Double DQN and Dueling DQN algorithms and optimising via the Prioritised Experience Replay mechanism and the Softmax action selection strategy, the PS-D3QN algorithm not only maintains swift convergence but also exhibits stability, effectively navigating potential local convergence issues while demonstrating notable learning and generalisation capabilities within the context of the current problem.

The PS-D3QN algorithm proposed in this study can also be applied to dynamic scheduling challenges in other fields, such as production and manufacturing, after reconfiguring the MDP model for specific problems with extensive discrete action and state spaces. It demonstrates commendable learning and generalisation capabilities in such scenarios.

4. Conclusion

This study presents a novel real-time scheduling approach called PS-D3QN based on the improved DQN algorithm. This method amalgamates the meritorious aspects of the Double DQN and Dueling DQN algorithms and employs two dueling neural networks. It adeptly gauges the Q-value of the target network by virtue of the action elected through the Q-value of the currently evaluating network. Moreover, the optimisation has been further finetuned by the ingenious design of a rational Prioritised Experience Replay (PER) mechanism and the integration of a Softmax action selection strategy. Additionally, this study examined the berth and yard scheduling predicaments prevalent in bulk cargo terminals. It has crafted an MDP model specifically for the scheduling issue, with the overarching goal of minimising both the cumulative time ships remain in the port and the total cost associated with cargo transportation. This was achieved by ingeniously amalgamating the authentic port milieu and pertinent data to configure the MDP model's state space, action space, and reward function.

Employing the PS-D3QN algorithm to address the scheduling conundrums based on actual ship, berth, and yard data yielded commendable optimization outcomes. In contrast with existing scheduling strategies and two alternative deep reinforcement learning algorithms, the PS-D3QN algorithm, as proposed in this study, has exhibited a substantial enhancement in the efficacy of berth and yard scheduling in port operations. Furthermore, it has contributed to the reduction of operational costs for ports while simultaneously mitigating the inherent empirical bias that arises from manual scheduling.

In the realm of future research endeavours, there exists the potential to elevate the algorithm's performance and stability through the refinement of the neural network architecture and the strategic selection of fitting optimisers. It is noteworthy that the intricacies of loading and unloading processes within bulk cargo ports involve a broader spectrum of resources. This study's scope was delimited to the berth and yard allocation facets of scheduling optimisation. The algorithm in this study holds the promise of expansion and applicability, potentially extending to more intricate challenges. By extending the model, PS-D3QN can address other resource scheduling problems at bulk cargo terminals, such as machinery and equipment scheduling. Furthermore, the algorithm can be applied to scheduling problems in production and manufacturing fields by re-establishing the MDP model, such as the job shop scheduling problem and its extensions.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (Grant No. 52172311) and China State Railway Group Co., Ltd. (Grant No. L2021X001).

References

[1] United Nations Conference on Trade and Development (2022). *Review of Maritime Transport 2022*, United Nations Publications, New York, USA, doi: 10.18356/9789210021470.

[2] Bouzekri, H., Alpan, G., Giard, V. (2023). Integrated laycan and berth allocation problem with ship stability and conveyor routing constraints in bulk ports, *Computers & Industrial Engineering*, Vol. 181, Article No. 109341, doi: 10.1016/j.cie.2023.109341.

[3] Bierwirth, C., Meisel, F. (2010). A survey of berth allocation and quay crane scheduling problems in container terminals, *European Journal of Operational Research*, Vol. 202, No. 3, 615-627, doi: 10.1016/j.ejor.2009.05.031.

[4] Bierwirth, C., Meisel, F. (2015). A follow-up survey of berth allocation and quay crane scheduling problems in container terminals, *European Journal of Operational Research*, Vol. 244, No. 3, 675-689, doi: 10.1016/j.ejor.2014.12.030.

[5] Đelović, D., Medenica Mitrović, D. (2017). Some considerations on berth productivity referred on dry bulk cargoes in a multipurpose seaport, *Tehnički Vjesnik – Technical Gazette,* Vol. 24, Supplement 2, 511-519, doi: 10.17559/TV-20150226074034.

[6] Barros, V.H., Costa, T.S., Oliveira, A.C.M., Lorena, L.A.N. (2011). Model and heuristic for berth allocation in tidal bulk ports with stock level constraints, *Computers & Industrial Engineering*, Vol. 60, No. 4, 606-613, doi: 10.1016/j.cie.2010.12.018.

[7] Umang, N., Bierlaire, M., Vacca, I. (2013). Exact and heuristic methods to solve the berth allocation problem in bulk ports, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 54, 14-31, doi: 10.1016/j.tre.2013.03.003.

[8] Ribeiro, G.M., Mauri, G.R., Beluco, S.d.C., Lorena, L.A.N., Laporte, G. (2016). Berth allocation in an ore terminal with demurrage, despatch and maintenance, *Computers & Industrial Engineering*, Vol. 96, 8-15, doi: 10.1016/j.cie.2016.03.005.

[9] Ernst, A.T., Oğuz, C., Singh, G., Taherkhani, G. (2017). Mathematical models for the berth allocation problem in dry bulk terminals, *Journal of Scheduling*, Vol. 20, No. 5, 459-473, doi: 10.1007/s10951-017-0510-8.

[10] Cheimanoff, N., Fontane, F., Kitri, M.N., Tchernev, N. (2021). A reduced VNS based approach for the dynamic continuous berth allocation problem in bulk terminals with tidal constraints, *Expert Systems with Applications*, Vol. 168, Article No. 114215, doi: 10.1016/j.eswa.2020.114215.

[11] Tang, L., Sun, D., Liu, J. (2016). Integrated storage space allocation and ship scheduling problem in bulk cargo terminals, *IIE Transactions*, Vol. 48, No. 5, 428-439, doi: 10.1080/0740817X.2015.1063791.

[12] Rocha de Paula, M., Boland, N., Ernst, A.T., Mendes, A., Savelsbergh, M. (2019). Throughput optimisation in a coal export system with multiple terminals and shared resources, *Computers & Industrial Engineering*, Vol. 134, 37-51, doi: 10.1016/j.cie.2019.05.021.

[13] Robenek, T., Umang, N., Bierlaire, M., Ropke, S. (2014). A branch-and-price algorithm to solve the integrated berth allocation and yard assignment problem in bulk ports, *European Journal of Operational Research*, Vol. 235, No. 2, 399-411, doi: 10.1016/j.ejor.2013.08.015.

[14] Unsal, O., Oguz, C. (2019). An exact algorithm for integrated planning of operations in dry bulk terminals, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 126, 103-121, doi: 10.1016/j.tre.2019.03.018.

[15] Tian, W., Zhang, H.P. (2021). A dynamic job-shop scheduling model based on deep learning, *Advances in Production Engineering & Management*, Vol. 16, No. 1, 23-36, doi: 10.14743/apem2021.1.382.

[16] François-Lavet, V., Henderson, P., Islam, R., Bellemare, M.G., Pineau, J. (2018). An introduction to deep reinforcement learning, *Foundations and Trends® in Machine Learning*, Vol. 11, No. 3-4, 219-354, doi: 10.1561/2200000071.

[17] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D. (2015). Human-level control through deep reinforcement learning, *Nature*, Vol. 518, 529-533, doi: 10.1038/nature14236.

[18] Van Hasselt, H., Guez, A., Silver, D. (2016). Deep reinforcement learning with double q-learning, In: *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence,* Palo Alto, California, USA, 12-17, doi: 10.1609/aaai.v30i1.10295.

[19] Wang, Z., Schaul, T., Hessel, M., Van Hasselt, H., Lanctot, M., de Freitas, N. (2016). Dueling network architectures for deep reinforcement learning, In: *Proceedings of the 33rd International Conference on Machine Learning*, Brookline, USA, 1995-2003, doi: 10.48550/arXiv.1511.06581.

[20] Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., Wierstra, D. (2015). Continuous control with deep reinforcement learning, *Machine Learning,* doi: 10.48550/arXiv.1509.02971.

[21] Sivamayil, K., Rajasekar, E., Aljafari, B., Nikolovski, S., Vairavasundaram, S., Vairavasundaram, I. (2023). A systematic study on reinforcement learning based applications, *Energies*, Vol. 16, No. 3, Article No. 1512, doi: 10.3390/en16031512.

[22] Lee, R.-Y., Chai, T.-Y., Chua, S.-Y., Lai, Y.-L., Wai, S.Y., Haw, S.-C. (2022). Cashierless checkout vision system for smart retail using deep learning, *Journal of System and Management Sciences*, Vol. 12, No. 4, 232-250, doi: 10.33168/JSMS.2022.0415.

[23] Kim, H.-J., Madhavi, S. (2022). A reinforcement learning model for quantum network data aggregation and analysis, *Journal of System and Management Sciences*, Vol. 12, No. 1, 283-293, doi: 10.33168/JSMS.2022.0120.

[24] Li, R., Zhang, X, Jiang, L, Yang, Z, Guo, W. (2022). An adaptive heuristic algorithm based on reinforcement learning for ship scheduling optimization problem, *Ocean & Coastal Management*, Vol. 230, Article No. 106375, doi: 10.1016/j.ocecoaman.2022.106375.

[25] Dai, Y., Li, Z., Wang, B. (2023). Optimizing berth allocation in maritime transportation with quay crane setup times using reinforcement learning, *Journal of Marine Science and Engineering*, Vol. 11, No. 5, Article No. 1025, doi: 10.3390/jmse11051025.

[26] Li, C., Wu, S., Li, Z., Zhang, Y., Zhang, L., Gomes, L. (2022). Intelligent scheduling method for bulk cargo terminal loading process based on deep reinforcement learning, *Electronics*, Vol. 11, No. 9, Article No. 1390, doi: 10.3390/electronics11091390.

[27] Hayes, C.F., Rădulescu, R., Bargiacchi, E., Källström, J., Macfarlane, M., Reymond, M., Verstraeten, T., Zintgraf, L.M., Dazeley, R., Heintz, F., Howley, E., Irissappane, A.A., Mannion, P., Nowé, A., Ramos, G., Restelli, M., Vamplew, P., Roijers, D.M. (2022). A practical guide to multi-objective reinforcement learning and planning, *Autonomous Agents and Multi-Agent Systems*, Vol. 36, No. 26, 1-59, doi: 10.1007/s10458-022-09552-y.

APEM journal

Advances in Production Engineering & Management
Volume 18 | Number 3 | September 2023 | pp 317–326
https://doi.org/10.14743/apem2023.3.475

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

Impact of agile, condition-based maintenance strategy on cost efficiency of production systems

Bányai, Á.[a,*]

[a]Institute of Logistics, University of Miskolc, Miskolc, Hungary

A B S T R A C T

Maintenance plays an increasingly important role in the life of production companies, as professional maintenance is an important prerequisite for the reliable operation of resources. A well-chosen maintenance strategy can make a major contribution to increased efficiency of production processes. The main goal of this research is to propose a novel optimization approach to define optimal maintenance strategy that ensures the efficient operation of the production process while reducing maintenance costs. The developed optimization method is based on Howard's policy iteration and describes the objective of the planning as a Markov decision process. The novelty and the scientific contribution of the presented study is the application of Howard's policy iteration methodology in a Markov decision process for agile, condition-based maintenance strategy optimization. As the results of the numerical analysis of the scenarios shows, the implementation of an optimized maintenance strategy based on the proposed approach can significantly increase the maintenance efficiency of the production process. The main reason for this is that the level and type of maintenance is always implemented depending on the current state of the system components, which reduces both the maintenance cost and the losses due to production downtime.

A R T I C L E   I N F O

Keywords:
Agile maintenance strategy;
Productivity;
Process control;
Markov decision process;
Maintenance strategy;
Optimization;
Smart manufacturing

**Corresponding author:*
agota.banyaine@uni-miskolc.hu
(Bányai, Á.)

Article history:
Received 9 May 2023
Revised 17 October 2023
Accepted 13 November 2023

1. Introduction

The global maintenance market is expected to grow from 42.66 billion USD in 2022 to 72.46 billion USD by 2029, and this grow means 7.9 % Compound Annual Growth Rate [1]. This fact shows the importance of maintenance in manufacturing. Across industry, a wide range of maintenance strategies can be used to support the availability of technological and logistics resources. These strategies can be classified in many ways. The maintenance strategies can be classified as preventive or corrective types. Preventive maintenance strategies are based on the idea, that maintenance operations are performed before failure occurs, while in the case of corrective maintenance, the maintenance operations are performed after the failure has occurred. However, Telek concluded in maintenance logistics research [2], that the maintenance appears as an independent service element of the production process, but in my opinion, maintenance strategy and maintenance operations must integrated into the whole business process, including purchasing, production, distribution and reverse processes. Agility can be a very important benefit of a well-chosen maintenance strategy, as it allows to react to detected failures in a timely manner through a well-chosen maintenance operation. A maintenance strategy can be considered well-chosen if it ensures efficient operation of the machines in a cost-effective way, so agile maintenance makes it possible to respond to changes in the condition level of the manufacturing plant.

A significant link between smart manufacturing and intelligent maintenance has been created by the fourth industrial revolution, which transformed conventional manufacturing systems into cyber-physical systems, creating real-time decision algorithms that can greatly increase the utilisation of production and logistics capacity in manufacturing systems, increasing flexibility and availability, while also greatly improving process sustainability. The connection between the smart manufacturing paradigm and the intelligent maintenance is based on digital twin technologies, which makes it possible to forecast future status of physical systems and make real-time decision regarding the maintenance strategy [3, 4].

As the literature review section shows, the existing research works are focusing on a wide range of optimization problems regarding maintenance, but only a few of them discuss the agile, condition-based maintenance. Based on this fact, the scope of this work is to propose a novel optimization approach to find a cost-efficient strategy for an agile, condition-based maintenance.

This paper is organised as follows. Section 2 presents a literature review focusing on the topic of maintenance policy optimization. Section 3 proposes a novel mathematical model, which makes it possible to define the cost-efficient strategy for agile, corrective maintenance. The section describes the transformation of the conventional P-F curve into a discrete P-F curve, which makes it possible to discretize the lead time from the possible detection point to the functional failure. The model can be described as a Markov decision process. Section 4 discusses the results of the numerical analysis of a scenario, which validates the mathematical model and the optimisation algorithm. Conclusions, future research directions and managerial impacts are discussed in Section 5.

2. Literature review

Within the frame of this section, I summarise the main results of maintenance strategy optimization related research results. I focus on the state-of-the-art technologies and give an overview of the most recent achievements in the field of maintenance strategy optimisation, in order to identify the bottlenecks that can be used to validate the research of agile, condition-based maintenance strategy optimization.

Li *et al* [3] in a multi-objective maintenance optimization concluded, that in the case of uncertain environment, it is possible, that the chosen maintenance strategy and performed maintenance operation is inappropriate, therefore integrated decision-making methodologies can be used to improve the conventional decision-making models to probabilistic uncertainty models. As Shi *et al*. [4] in a research work focusing on preventive maintenance strategy found, it is important to focus on the lifecycle safety and availability of the maintained system, which applies especially to the preventive maintenance strategies, where the decomposition of lifecycle failure states and lifecycle failure probability plays also an important role in the modelling of the optimised maintenance strategy. The integration of inspection and maintenance is a suitable, but challenging improvement direction of maintenance strategies. Guo and Liang [5] concluded in a study describing the optimization of maintenance strategies as predictive Markov decisions, that inspection and maintenance strategies must be flexible, because in the early phase of the lifecycle of the inspected and maintained system, predictive inspections are not needed as often as in later phase of the lifecycle of the system, which can lead to wastes of human and technological resources. The lifecycle related problems are also discussed by Hernández *et al*. [6], and their approach shows, that the maintenance of networked assets with progressively deteriorating condition levels can also be optimized considering the dynamics of data traffic.

Zhang *et al*. [7] found in a study regarding emergency maintenance, that the optimization of maintenance strategies is particularly complex when the scheduling of maintenance operations needs to be integrated with the scheduling of the operation of the system being maintained, a task that can become particularly complex for a hyper-connected complex system such as a high-speed railway lines. In this case, the rolling horizon framework is a suitable tool to perform real-time implementation of decisions and maintenance operations. Pinciroli *et al*. [8] discusses a same topic, focusing on the integrated optimization of operation and maintenance of renewable energy systems. The reliability-centered maintenance (RCM) is suggested by Paoprasert *et al*. [9] to improve key performance indicators of a HDD production system. As the study concluded, RCM is suitable to increase availability of machines and reliability of the manufacturing system. However, Industry 4.0 focuses on technology, but a survey on enabling technologies [10] shows, that the upcoming industrial revolution will be directed to the operators, which means, that the role of human resources in maintenance systems will continue to grow, despite increasing technological support, which will lead to new challenges, in particular for the training of human resources. Case studies validated by Mappas *et al*. [11], that automated maintenance operations can significantly increase the efficiency of maintenance operations. Based on these studies, we can conclude, that maintenance strategy optimization is an extensively researched topic, including a wide range of models, solution algorithms and application fields.

The used models and methods include the followings: Monte Carlo method [3], stochastic simulation [4], Forward algorithm [5], Baum-Welch algorithm [5], Lagrangian relaxation [7], mixed-integer nonlinear optimization [7], artificial intelligence [10], deep reinforcement learning [8] and Failure-Mode-and-Effect-Analysis (FMEA) [12], Fuzzy-TOPSYS [13], simulated annealing [14] and other heuristics [15]. The applications and case studies includes a wide range of industries: wind farms [3], high-speed railway [7], HDD manufacturing [9], automotive [12], injection moulding [16], offshore floating systems [17] and nuclear power plant [18] and they analyse different types of maintenance solutions including preventive [4], predictive [5, 19], emergency [7], condition-based [12] and collaborative maintenance [20], and measurement of maintenance excellence from technical and financial point of view [21].

The consequences of the literature review are the followings:

- The articles that addressed the optimization of maintenance strategies are focusing on different maintenance types, but only a few of them discusses the agile, condition-based maintenance.
- A wide range of research articles discuss the optimization of maintenance strategies using the conventional P-F curve [22], but the transformation of this continuous P-F curve into a discretised P-F curve to describe transition probabilities between different condition levels of machines and plant is not discussed as a potential tool to integrate the cost efficiency of both manufacturing and maintenance. Therefore, this research topic still needs more attention and research.
- Mathematical models and solution algorithms are important tools for the optimization of maintenance strategies, which can lead to increased quality [23]. According to that, the main goal of this research is to propose a novel mathematical model and solution algorithm to support the optimization of agile condition-based maintenance.

3. Materials and methods

Developing an optimal strategy for maintenance processes can be defined as an assignment problem, where maintenance operations of different types, depth and cost are assigned to the technological and logistics resources, in order to ensure the efficient, smooth operation of the production process and improve the availability of machines and plant. The following assumptions can be used in this assignment task. We can define different conditions of the technological and logistics system, which can be monitored accurately in real-time either by a digital twin solution or by a conventional sensor-based monitoring of technological and logistics resources as a digital shadow of the real-world system:

$$C = (c_1, c_2, \cdots, c_i, \cdots, c_\alpha) \tag{1}$$

where c_i is the condition level *i* of the system and state *i* of the system and α defines the potential conditional levels depending on the condition levels of technological and logistics resources. Condition monitoring makes it possible to collect information regarding the condition of the technological and logistics resources including temperature, pressure, vibration, abrasion, noise. The condition level of the system significantly influences the availability of the machines and plant, because low condition level can lead to downtime or increased reject rate (lower product quality).

The transition between these condition levels can occur for two reasons. One is when the condition of the system decreases during continuous operation, causing the system's condition level to decrease. The other is when the condition of the machines and the plant improves because of a maintenance or condition improvement operation. We can also define a set of potential maintenance operations (or maintenance levels) which can significantly influence the transition between two potential condition levels.

$$M^q = (m_1^q, m_2^q, \cdots, m_j^q, \cdots, m_\beta^q) \tag{2}$$

where m_j^q is the potential maintenance operation j for maintenance strategy q, β is the upper limit of potential maintenance operations. We can define the upper limit of the maintenance operations as a dynamic parameter, because depending on the new, unknown condition levels, new maintenance operations can be defined and set up. It means, that $\alpha = \alpha(t)$ and $\alpha = \alpha(t)$.

The above-mentioned transition probabilities can statistically describe the probability between two predefined condition levels. For example, if a drilling machine is working properly (condition level 1) but the temperature of the drilling tool exceeds 175° C (condition level 2) than it can lead to decreased product quality (condition level 3) and it can also lead to failure in product (condition level 4) and machine (condition level 5). The transition probabilities

The transition probabilities define the basis for the selection of the optimal maintenance operation, as different maintenance operations lead to different condition levels.

$$\forall\gamma: \sum_{i=1}^{\alpha} t_{i\gamma} = 1 \tag{3}$$

where $T = [t_{i\gamma}]$ is the transition probability matrix defining the transition probability between condition level i and condition level γ.

In conventional condition-based maintenance models, the P-F curve describes that as a failure starts manifesting, the machine or plant deteriorates to the point at which it can possibly be detected (point P). If the failure is not detected, it continues until a functional failure occurs (point F) [23]. Using transition probabilities of potential condition levels of machines and plant, it is possible to transform this conventional P-F curve into a discretised P-F curve, as Fig. 1 shows.

We can assign cost to both to these condition levels, reflecting machine and system availability, productivity, and product quality and to the maintenance operations.

It is important to note, that the elements of the transition matrix are highly influenced by maintenance operations. If maintenance operation δ is performed in the case of condition level i of the system, then the transition probability from condition level i to condition level j of the system is not necessarily the same as the transition between condition level i to condition level j of the system. The reason for this is, that the maintenance operation results a condition level improvement from condition level i to a new condition level k, and the probability of a transition from condition level i to condition level j is therefore depends on the transition between condition level k and condition level j.

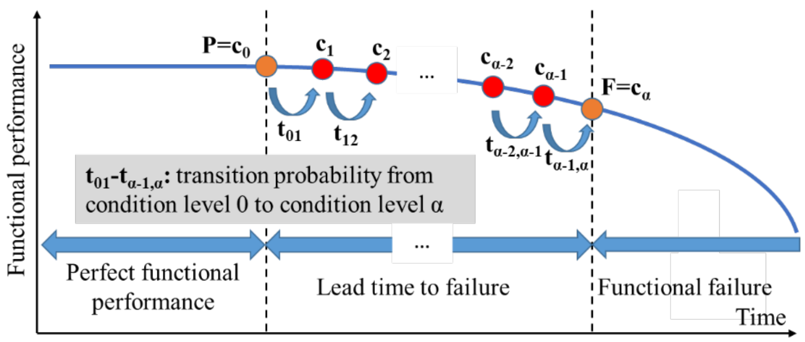


Fig. 1 Discretized P-F curve describing transition probabilities between condition levels

$$p(c_j|c_i, m_\delta) \in \mathrm{T} \wedge p(c_j|c_i, m_\delta) \leq t_{ij} \vee p(c_j|c_i, m_\delta) \geq t_{ij} \tag{4}$$

When planning maintenance processes, we can use different objective functions to find the optimal solution. In my previous study [24], I have shown the energy efficiency-based maintenance policy optimization. In this model, the discounted profit based on the maintenance cost and cost of lost production is the objective function. The optimization problem is a Markov decision problem; therefore, it is also an infinite horizon probabilistic dynamic programming problem, and the objective function is the optimization of the discounted profit as follows:

$$DP^0(c_i) = pe_{c_i, m_\delta(c_i)} + \varepsilon \cdot \sum_{\delta=1}^{\beta} p\left(c_j \middle| c_i, m_\delta(c_i)\right) DP^0(c_j) \tag{5}$$

where: $DP^0(c_i)$ is the expected discounted profit depending on the condition level of the manufacturing system, $pe_{c_i, m_\delta(c_i)}$ is the expected profit depending on the chosen maintenance operation m_δ and condition level c_i, ε is the discounting factor, which can significantly influence the value of the objective function, because higher discounting factor lead to higher discounted profit.

The expected profit can be defined in many ways. In this approach, the expected profit is defined depending on the following parameters: expected income resulted by MRP (Materials Requirement Planning), lost value caused by the downtime and cost of maintenance operations:

$$\mathrm{H}_M^0(c_i) = \max_{m_\delta \in M(c_i)} \left(pe_{c_i, m_\delta(c_i)} + \varepsilon \cdot \sum_{\delta=1}^{\delta_{max}} p\left(c_j \middle| c_i, m_\delta(c_i)\right) DP^0(c_j) \right) \tag{6}$$

where $\mathrm{H}_M^0(c_i)$ is the Howard's parameter in the case of condition level *i* of the manufacturing system. Based on Eqs. 5 and 6 we can compare the Howard's parameter and the discounted profit value. If the Howard's parameter is equal to the discounted profit, then M is the optimal maintenance strategy. Otherwise, the maintenance operations assigned to each condition level must be changed and then both parameters must be recomputed.

4. Results and discussion

Within the frame of the scenario analysis, a U-shaped manufacturing system is analysed including 10 machines. 10 different conditions levels of the manufacturing system are defined: $C = (c_1, \cdots, c_{10})$, where c_1 represents the best condition level, as Table 1 shows.

Table 1 Transition probabilities of condition levels in the manufacturing system

$T = [t_{i\gamma}]$	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
c_1	0.6	0.25	0.1	0.05	0	0	0	0	0	0
c_2	0	0.5	0.3	0.1	0.1	0	0	0	0	0
c_3	0	0	0.82	0.11	0.05	0.02	0	0	0	0
c_4	0	0	0	0.4	0.3	0.15	0.07	0.05	0.03	0
c_5	0	0	0	0	0.7	0.2	0.1	0	0	0
c_6	0	0	0	0	0	0.5	0.4	0.1	0	0
c_7	0	0	0	0	0	0	0.7	0.2	0.1	0
c_8	0	0	0	0	0	0	0	0.6	0.4	0
c_9	0	0	0	0	0	0	0	0	0.8	0.2
c_{10}	0	0	0	0	0	0	0	0	0	1

There are 10 potential maintenance operations in this scenario, which can be performed depending on the current condition level of the manufacturing system. The OMS (online monitoring system) makes it possible to collect data regarding condition level of the machines and plant and defines the expected transition possibilities between condition.

Maintenance operations are assigned to each condition level of the manufacturing system. The probability that the production system will move from one condition level to another can be determined by computing the condition level resulted from the maintenance operation and after that we can calculate the transition probability. As an example, in the case of transition probability t_{35}, we can calculate the potential values as:

$$p(c_5|c_3, m_0) = p(c_5|c_4, m_1) = \cdots = p(c_9|c_9, m_6) = p(c_9|c_{10}, m_7) = t_{35} = 0.05 \tag{7}$$

and this calculation of transition probabilities can be generalized as follows:

$$p\forall\ \rho-\sigma=\gamma \wedge t_{i\gamma}>0: p\left(c_\gamma \middle| c_\rho, m_\sigma\right)=t_{i\gamma} \tag{8}$$

Let the expected income resulted by the material requirement planning be $c(MRP) = 30000$ €. The lost value of the manufacturing process, depending on the condition level of the manufacturing system can be also defined as follows as shown in Table 2:

$$lv_{c_i} = ilv^{\varphi^{c_i}} \tag{9}$$

where ilv is the initial lost value, which is in this scenario 12000 €, φ is the specific parameter influencing the lost value depending on the condition level of the manufacturing system.

We can define in the same way the maintenance cost depending on the condition level of the manufacturing system and the performed maintenance operation as follows:

$$mc_{c_i,m_\delta(c_i)} = imc^{\omega^{m_\delta}} \tag{10}$$

where imc is the initial maintenance cost, which is in this scenario 2500 €, ω is the specific parameter influencing the maintenance cost depending on the condition level of the manufacturing system and the performed maintenance operation. The computed values of the scenario analysis are shown in Table 3.

As Table 3 shows, the maintenance operation have increased cost depending of the complexity of them, because complex maintenance operations can lead to a more significant condition level improvement of the machines and plant.

Based on Eqs. 6, 9, and 10 we can compute the expected profit of the scenario depending on the current condition level of the machines and plant and the assigned maintenance operation, as shown in Table 4. Let define the initial maintenance strategy by the assignment of maintenance operations to condition levels as given:

$$A^0 = \left[a_{c_i}^0\right] = [m_0, m_1, m_1, m_2, m_1, m_3, m_2, m_1, m_4, m_2] \tag{11}$$

where $a_{c_i}^0 = m_\pi$, and is the maintenance operation π is assigned to condition level c_i in the initial phase of the optimization.

Table 2 Lost value of the manufacturing system depending on the condition level in (€)

c_i	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
lv_{c_i}	18000	17185	16307	15363	14347	13252	12072	10800	9427	7946

Table 3 Maintenance cost depending on the condition level and maintenance operation in (€)

m_i	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9
$mc_{c_i,m_\delta(c_i)}$	2500	4734	6641	9443	13618	19929	29611	44701	68601

Table 4 Expected profit of the scenario in (€)

$pe_{c_i,m_\delta(c_i)}$	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
m_0	18000	17185	16307	15363	14347	13252	12072	10800	9427	7946
m_1	-	15500	14685	13807	12863	11847	10752	9572	8300	6927
m_2	-	-	13266	12451	11573	10629	9613	8518	7338	6066
m_3	-	-	-	11359	10544	9666	8722	7706	6611	5431
m_4	-	-	-	-	8557	7742	6864	5920	4904	3809
m_5	-	-	-	-	-	4382	3566	2689	1745	729
m_6	-	-	-	-	-	-	-1929	-2744	-3621	-4565
m_7	-	-	-	-	-	-	-	-11611	-12427	-13304
m_8	-	-	-	-	-	-	-	-	-26701	-27516
m_9	-	-	-	-	-	-	-	-	-	-50601

Once the input parameters for the scenario have been defined, the discounted profit of the initial maintenance strategy can be computed based on Eq. 5 solving the following value definition equations:

$$DP^0(c_1) = pe_{c_1,m_0} + \varepsilon \cdot \sum_{\theta=1}^{4} t_{1\theta} \cdot DP^0(c_\theta) \tag{12}$$
$$DP^0(c_2) = pe_{c_2,m_1} + \varepsilon \cdot \sum_{\theta=1}^{4} t_{1\theta} \cdot DP^0(c_\theta) \tag{13}$$
$$DP^0(c_3) = pe_{c_3,m_1} + \varepsilon \cdot \sum_{\theta=2}^{5} t_{2\theta} \cdot DP^0(c_\theta) \tag{14}$$
$$DP^0(c_4) = pe_{c_4,m_2} + \varepsilon \cdot \sum_{\theta=2}^{5} t_{2\theta} \cdot DP^0(c_\theta) \tag{15}$$
$$DP^0(c_5) = pe_{c_5,m_1} + \varepsilon \cdot \sum_{\theta=4}^{9} t_{4\theta} \cdot DP^0(c_\theta) \tag{16}$$
$$DP^0(c_6) = pe_{c_6,m_3} + \varepsilon \cdot \sum_{\theta=3}^{6} t_{3\theta} \cdot DP^0(c_\theta) \tag{17}$$
$$DP^0(c_7) = pe_{c_7,m_2} + \varepsilon \cdot \sum_{\theta=5}^{7} t_{5\theta} \cdot DP^0(c_\theta) \tag{18}$$
$$DP^0(c_8) = pe_{c_8,m_1} + \varepsilon \cdot \sum_{\theta=7}^{9} t_{7\theta} \cdot DP^0(c_\theta) \tag{19}$$
$$DP^0(c_9) = pe_{c_9,m_4} + \varepsilon \cdot \sum_{\theta=5}^{7} t_{5\theta} \cdot DP^0(c_\theta) \tag{20}$$
$$DP^0(c_{10}) = pe_{c_{10},m_2} + \varepsilon \cdot \sum_{\theta=8}^{9} t_{8\theta} \cdot DP^0(c_\theta) \tag{21}$$

The solution of the above-mentioned value definition equations resulted the discounted profit for the initial maintenance strategy describing the assignment of maintenance strategies to condition levels of the machines and plant shown in Table 5.

As Table 5 shows, the condition level of the machines and plant has a significant impact on the discounted value, because lower condition levels lead to lower discounted value.

Table 5 Discounted profit for the initial maintenance strategy in (€)

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
$DP^0(c_i)$	324528	322028	316994	314760	306318	309936	300776	293474	296067	285852

The next phase of the optimization is to check the validity of the initial maintenance strategy. Based on Eq. 6, it is possible to calculate the Howard's parameter for each condition level of the machines and plant, and then modify the initial maintenance strategy based on the maximum value of the Howard's parameter. In the case of c_1 condition level, the Howard's parameter is the same the discounted value, therefore no maintenance strategy modification is required in the second iteration phase.

$$\mathrm{H}_M^0(c_1) = DP^0(c_1) \rightarrow a_{c_1}^1 = a_{c_1}^0 = m_0 \tag{22}$$

In the case of c_2 condition level, we can calculate the Howard's parameter based on Eq. 6, as follows:

$$\mathrm{H}_M^0(c_2) = max \left| \begin{matrix} m_0 \rightarrow pe_{c_2,m_0} + \varepsilon \cdot \sum_{\theta=2}^{5} t_{2\theta} \cdot DP^0(c_\theta) \\ m_1 \rightarrow pe_{c_2,m_1} + \varepsilon \cdot \sum_{\theta=1}^{4} t_{1\theta} \cdot DP^0(c_\theta) \end{matrix} \right. \tag{23}$$

The comparison of the Howard's parameter and the discounted value shows, that no maintenance strategy change is required in the case of condition level c_2.

$$\mathrm{H}_M^0(c_2) = DP^0(c_2) \rightarrow a_{c_2}^1 = a_{c_2}^0 = m_1 \tag{24}$$

In the case of c_3 condition level, we can calculate the Howard's parameter in the same way:

$$\mathrm{H}_M^0(c_3) = max \left| \begin{matrix} m_0 \rightarrow pe_{c_3,m_0} + \varepsilon \cdot \sum_{\theta=3}^{6} t_{3\theta} \cdot DP^0(c_\theta) \\ m_1 \rightarrow pe_{c_3,m_1} + \varepsilon \cdot \sum_{\theta=2}^{5} t_{3\theta} \cdot DP^0(c_\theta) \\ m_2 \rightarrow pe_{c_3,m_2} + \varepsilon \cdot \sum_{\theta=1}^{4} t_{3\theta} \cdot DP^0(c_\theta) \end{matrix} \right. \tag{25}$$

The comparison of the Howard's parameter and the discounted value shows, that we can change maintenance operation m_1 to maintenance operation m_2 assigned to condition level c_3.

$$\mathrm{H}_M^0(c_3) > DP^0(c_3) \rightarrow a_{c_3}^1 \neq a_{c_3}^0 \rightarrow a_{c_3}^1 = m_2 \tag{26}$$

This maintenance operation change resulted a 2801 € additional discounted profit in the case of condition level c_3. We can calculate the new maintenance operations assigned to each condition level leading to increased discounted profit in the same way. As Table 6 shows, the iterative methodology after the first iteration phase lead to the change of 8 assignment of maintenance operations to condition levels, and the value of the total additional discounted value can be calculated as follows:

$$\forall z > 0: TADP^z = \sum_{i=1}^{\alpha} ADP^z(c_i) = \sum_{i=1}^{\alpha} DP^z(c_i) - DP^{z-0}(c_i) = 39446\ € \quad (27)$$

where $TADP^z$ is the total additional discounted value after iteration phase z, $ADP^z(c_i)$ is the additional discounted value after iteration phase z in the case of condition level i.

The above-described iterative calculation process must be continued as long as it is possible to increase the total discounted value by changing the maintenance strategy. The final result of the maintenance strategy optimisation is shown in Table 7.

Table 6 The increased discounted value per condition level in (€) and the assignment of maintenance strategies and condition levels after the first iteration

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
$a_{c_i}^1$	m_0	m_1	m_2	m_3	m_4	m_5	m_6	m_5	m_6	m_4
$ADP^1(c_i)$	0	0	2801	3128	8768	975	6358	9484	581	7352

Table 7 The increased discounted value per condition level in (€) and the assignment of maintenance strategies and condition levels after the first iteration

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
$a_{c_i}^3$	m_0	m_1	m_2	m_3	m_4	m_5	m_6	m_5	m_4	m_9
$ADP^3(c_i)$	8294	8294	11095	11421	17061	10823	17180	20307	14258	20298

Based on the above discussed methodology, it can be concluded that the optimization of the maintenance strategy and the modification of the assignment of maintenance operations to machines can significantly contribute to the increase of the efficiency of the production system, since an overall increase of 130030 € in discounted value was achieved.

For the optimization method presented above, it is important to perform a sensitivity analysis of the objective function for some parameters. By analysing the impact of the maintenance cost and the discount rate on the discounted value, it can be concluded that an increase in the maintenance cost decreases the discounted value obtained by the maintenance strategy. This finding seems trivial, but the impact of the maintenance cost on the discounted value is not trivial, since a given increase in the maintenance cost does not change the discounted value resulting from the strategy to the same extent, since the productivity resulting from the condition level of the machines can be modelled as a Markov process with transition probabilities, as illustrated in Fig. 2.

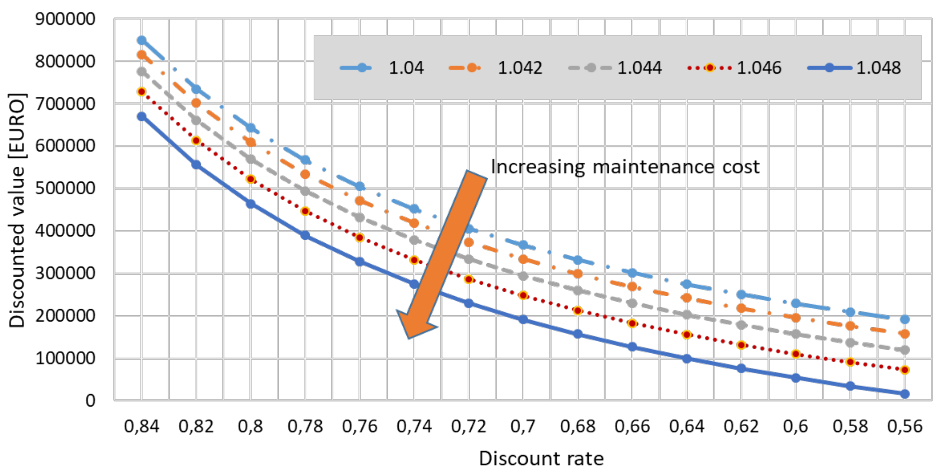


Fig. 2 Impact of maintenance cost and discount rate on discounted value of agile maintenance strategy

Fig. 2 also shows how a decrease in the discount rate has a decreasing effect on the discounted value, a relationship that can also be explained by the transition probabilities between the condition levels of the machines. Based on this line of thinking, it can be seen that the maintenance costs influenced by the maintenance operations, the revenues associated with the condition levels of machines and plant and other system parameters have a significant impact on the discounted value that can be achieved by implementing an agile maintenance strategy.

5. Conclusion

Optimising maintenance processes is an increasingly important goal of production companies. This is because, in order to meet the dynamically changing customer demands in a cost-effective way, the machines in the production system must be available at all times in a condition to produce a product of the right quality. In this paper, the author presents a maintenance strategy optimization methodology that is suitable for modelling the transitions between the condition levels of the manufacturing system as Markov chains and is suitable for efficient application of Markov decision process to select the optimal maintenance strategy. The presented iteration-based methodology is suitable for determining the optimal maintenance strategy. The essence of this methodology is that for each condition level the optimal maintenance operation can be determined, as a result of which the discounted value associated with that condition level can be increased, i.e. the lost value caused by downtime or rejected products can be reduced.

A new methodology has been developed in this research work and its applicability has been validated by the numerical analysis of a case study. After validation of the methodology, the following conclusions can be drawn:

- Optimisation of the maintenance strategy can significantly improve the efficiency of production systems, thereby enhancing product quality. The discounted value as a metric can be well used to measure this improvement.
- The presented iterative method is not only suitable for the optimization of small-scale tasks but is also well suited for multi-machine manufacturing systems, since the presented methodology does not include complex computational procedures that would be computationally expensive, i.e. the presented optimization task does not belong to the NP-hard optimization problems.
- The maintenance strategy is optimized in iterative steps. In each iteration phase, the discounted value is gradually increased.

The presented new methodology is important not only for the academy, but its practical applicability can also be significant since it can greatly contribute to the enhancement of the competitiveness of production companies by increasing the efficiency and availability of the manufacturing systems.

The research presented also has implications for managerial decisions, as the optimisation of the maintenance strategy can significantly influence the optimization of technological, logistics and human resources, and decisions can be taken on the outsourcing of certain processes (for example, the decision to outsource maintenance processes can be justified). The application of machine learning solutions can be defined as a potential future research direction. Cooperation and networking make it possible to improve the efficiency of maintenance operations [27], therefore the second potential future research direction is the development of novel optimization methods for networking companies.

References

[1] Fortune Business Inside. Inspection, repair and maintenance market, from *https://www.fortunebusinessinsights.com/inspection-repair-and-maintenance-market-102983*, accessed May 2, 2023.

[2] Telek, P. (2023). Logistics activities of maintenance processes, *Advanced Logistic Systems - Theory and Practice*, Vol. 17, No. 1, 49-54, doi: 10.32971/als.2023.006.

[3] Gao, G., Zhou, D., Tang, H., Hu, X. (2021). An intelligent health diagnosis and maintenance decision-making approach in smart manufacturing, *Reliability Engineering and System Safety*, Vol. 216, Article No. 107965, doi: 10.1016/j.ress.2021.107965.

[4] Cortés-Leal, A., Cárdenas, C., Del-Valle-Soto, C. (2022). Maintenance 5.0: Towards a worker-in-the-loop framework for resilient smart manufacturing, *Applied Sciences*, Vol. 12, No. 22, Article No. 11330, doi: 10.3390/app122211330.

[5] Li, M., Jiang, X., Carroll, J., Negenborn, R.R. (2022) A multi-objective maintenance strategy optimization framework for offshore wind farms considering uncertainty, *Applied Energy*, Vol. 321, Article No. 119284, doi: 10.1016/j.apenergy.2022.119284.
[6] Shi, Y., Lu, Z., Huang, H., Liu, Y., Li, Y., Zio, E., Zhou, Y. (2022). A new preventive maintenance strategy optimization model considering lifecycle safety, *Reliability Engineering & System Safety*, Vol. 221, Article No. 108325, doi: 10.1016/j.ress.2022.108325.
[7] Guo, C., Liang, Z. (2022). A predictive Markov decision process for optimizing inspection and maintenance strategies of partially observable multi-state systems, *Reliability Engineering & System Safety*, Vol. 226, Article No. 108683, doi: 10.1016/j.ress.2022.108683.
[8] Pérez Hernández, M. Puchkova, A., Parlikad, A.K. (2022). Maintenance strategies for networked assets, *IFAC-PapersOnLine*, Vol. 55, No. 19, 151-156, doi: 10.1016/j.ifacol.2022.09.199.
[9] Zhang, H., Li, S., Wang, Y., Yang, L., Gao, Z. (2021). Collaborative real-time optimization strategy for train rescheduling and track emergency maintenance of high-speed railway: A Lagrangian relaxation-based decomposition algorithm, *Omega*, Vol. 102, Article No. 102371, doi: 10.1016/j.omega.2020.102371.
[10] Pinciroli, L., Baraldi, P., Ballabio, G., Compare, M., Zio, E. (2022). Optimization of the operation and maintenance of renewable energy systems by deep reinforcement learning, *Renewable Energy*, Vol. 183, 752-763, doi: 10.1016/j.renene.2021.11.052.
[11] Paoprasert, N., Lin, W.Y.H., Muneekaew, T. (2022). Assessing risk priority numbers of failures in the screw tightening machine of a hard disk drive production system, *Journal of Machine Engineering*, Vol. 22, No. 1, 124-137, doi: 10.36897/jme/145272.
[12] Mourtzis, D., Angelopoulos, J., Panopoulos, N. (2022). Operator 5.0: A survey on enabling technologies and a framework for digital manufacturing based on extended reality, *Journal of Machine Engineering*, Vol. 22, No. 1, 43-69, doi: 10.36897/jme/147160.
[13] Mappas, V., Vassiliadis, V.S., Dorneanu, B., Routh, A.F., Arellano-Garcia, H. (2022). Maintenance scheduling optimisation of Reverse Osmosis Networks (RONs) via a multistage Optimal Control reformulation, *Desalination*, Vol. 543, Article No. 116105, doi: 10.1016/j.desal.2022.116105.
[14] Ramere, M.D., Laseinde, O.T. (2021). Optimization of condition-based maintenance strategy prediction for aging automotive industrial equipment using FMEA, *Procedia Computer Science,* Vol. 180, 229-238, doi: 10.1016/j.procs.2021.01.160.
[15] Patalas-Maliszewska, J., Losyk, H. (2022). An approach to maintenance sustainability level assessment integrated with Industry 4.0 technologies using Fuzzy-TOPSIS: A real case study*, Advances in Production Engineering & Management,* Vol. 17, No. 4, 455-468, doi: 10.14743/apem2022.4.448.
[16] Diaz Cazanas, R., Delgado Sobrino, D.R., Caganova, D., Kostal, P., Velisek, K. (2019). Joint programming of production-maintenance tasks: A Simulated Annealing-based method, *International Journal of Simulation Modelling,* Vol. 18, No. 4, 666-677, doi: 10.2507/IJSIMM18(4)503.
[17] Xu, E.B., Yang, M.S., Li, Y., Gao, X.Q., Wang, Z.Y., Ren, L.J. (2021). A multi-objective selective maintenance optimization method for series-parallel systems using NSGA-III and NSGA-II evolutionary algorithms, *Advances in Production Engineering & Management,* Vol. 16, No. 3, 372-384, doi: 10.14743/apem2021.3.407.
[18] Bleicher, F., Ramsauer, C., Leonhartsberger, M., Lamprecht, M., Stadler, P., Strasser, D., Wiedermann, C. (2021). Tooling systems with integrated sensors enabling data based process optimization, *Journal of Machine Engineering,* Vol. 21, No. 1, 5-21, doi: 10.36897/jme/134244.
[19] George, B., Loo, J., Jie, W. (2023). Novel multi-objective optimisation for maintenance activities of floating production storage and offloading facilities, *Applied Ocean Research,* Vol. 130, Article No. 103440, doi: 10.1016/j.apor.2022.103440.
[20] Peng, H., Wang, Y., Zhang, X., Hu, Q., Xu, B. (2022). Optimization of preventive maintenance of nuclear safety-class DCS based on reliability modeling, *Nuclear Engineering and Technology*, Vol. 54, No. 10, 3595-3603, doi: 10.1016/j.net.2022.05.011.
[21] Farahani, A., Tohidi, H., Shoja, A. (2019). An integrated optimization of quality control chart parameters and preventive maintenance using Markov chain, *Advances in Production Engineering & Management,* Vol. 14, No. 1, 5-14, doi: 10.14743/apem2019.1.307.
[22] Tavakoli Kafiabad, S., Zanjani, M.K., Nourelfath, M. (2022). Robust collaborative maintenance logistics network design and planning, *International Journal of Production Economics*, Vol. 244, Article No. 108370, doi: 10.1016/j.ijpe.2021.108370.
[23] Djurović, D., Bulatović, M., Soković, M., Stoić, A. (2015). Measurement of maintenance excellence, *Tehnički Vjesnik – Technical Gazette,* Vol. 22, No. 5, 1263-1268, doi: 10.17559/TV-20140922094945.
[24] Orošnjak, M., Brkljač, N., Šević, D., Čavić, M., Oros, D., Penčić, M. (2023). From predictive to energy-based maintenance paradigm: Achieving cleaner production through functional-productiveness, *Journal of Cleaner Production*, Vol. 408, Article No. 137177, doi: 10.1016/j.jclepro.2023.137177.
[25] Hupje, E. 9 Types of maintenance: How to choose the right maintenance strategy, from *https://roadtoreliability.com/types-of-maintenance/#h-condition-based-maintenance-cbm*, accessed April 22, 2023.
[26] Bányai, Á. (2021). Energy consumption-based maintenance policy optimization, *Energies,* Vol. 14, No. 18, Article No. 5674, doi: 10.3390/en14185674.
[27] Bányai, T., Veres, P., Illés, B. (2015). Heuristic supply chain optimization of networked maintenance companies, *Procedia Engineering,* Vol. 100, 46-55, doi: 10.1016/j.proeng.2015.01.341.

Advances in Production Engineering & Management **ISSN 1854-6250**
Volume 18 | Number 3 | September 2023 | pp 327–344 Journal home: apem-journal.org
https://doi.org/10.14743/apem2023.3.476 Original scientific paper

A game theory analysis of intelligent transformation and sales mode choice of the logistics service provider

Cao, G.M.[a,b]**, Zhao, X.X.**[c]**, Gao, H.H.**[d,*]**, Tang, M.C.**[e]

[a]College of Business Administration, Henan Finance University, Henan Province, P.R. China
[b]School of Economics and Management, Beijing Jiaotong University, Beijing, P.R. China
[c]Shijiazhuang Posts and Telecommunications Technical College, Hebei Province, P.R. China
[d]School of Management Engineering, Zhengzhou University of Aeronautics, Henan Province, P.R. China
[e]International Center for Informatics Research, Beijing Jiaotong University, Beijing, P.R. China

A B S T R A C T

In order to study whether the logistics service provider (LSP) should carry out intelligent transformation strategy of logistics services, this paper constructs a logistics service supply chain consisting of one LSP and one logistics service integrator (LSI), and discusses whether the LSP is independent or participate in LSI. The paper shows that choosing the intelligent transformation of logistics services under any mode can improve the profits of the LSP and the LSI. The joint transformation of logistics services to improve the profit of the LSI is not affected by the choice of mode, while the profit of LSP under the resale mode remains unchanged when she chooses joint intelligent transformation. When the intelligent transformation level is high, the LSI tends to choose the resale model; otherwise, the LSI tends to choose the platform model. When the LSP chooses intelligent transformation by herself, if the share ratio is low, the LSI tends to choose the resale model. If the share ratio is high and the level of intelligent transformation of logistics services is not high, the LSI more inclines to choose the platform model.

A R T I C L E I N F O

Keywords:
Logistics service supply chain (LSSC);
Intelligent transformation;
Sales model;
Decision analysis;
Logistics service provider (LSP);
Logistics service integrator (LSI);
Profit;
Game theory

**Corresponding author:*
wangji946@163.com
(Gao, H.H.)

Article history:
Received 3 December 2022
Revised 9 September 2023
Accepted 13 September 2023

1. Introduction

With the growing specialization, customization, and diversification of customer needs, the large-scale customized logistics service model emerges as the times require [1]. Under this kind of logistics service mode, logistics enterprises spontaneously form a LSSC logistics service supply chain through alliance and integration, and use the scale effect to reduce logistics service costs and meet the increasingly rich service needs of customers as much as possible under certain cost constraints [2, 3]. According to different functions, the LSSC mainly includes LSPs that provide basic logistics functions such as transportation, packaging, and warehousing, LSIs that integrate various basic logistics services, and final logistics service demand customers [3, 4]. Numerous studies have shown that under the background of mass customization, it is very effective to achieve specific goals of logistics service needs through LSIs in the LSSC and customers, and to use the basic functions of LSPs to finally meet customers' logistics service requirements [3, 5, 6].

The efficiency of the logistics service system can effectively solve the different types of service needs of customers and meet the challenges brought by changes in demand. In the operation process of enterprises, to improve the operational efficiency of logistics services, enterprises try to adopt emerging technologies such as cloud computing, the Internet of Things, big data, block-chain, etc. [7], the traditional mode of logistics has been subverted, and intelligent logistics has entered the era of logistics 4.0 under the background of Industry 4.0 [8, 9]. Based on traditional logistics, intelligent logistics uses modern intelligent information technology and intelligent equipment to identify and perceive all aspects of the logistics system, to achieve efficient control of the entire logistics system, and to make decisions that better meet customer needs [10]. Intelligent logistics has been widely used in many enterprises and achieved good results. For example, DHL integrates the Internet and autonomous driving technology in the logistics business and is responsible for the dedicated innovation center [11], JD Logistics is already experimenting with package delivery via drones in Xi'an [12], Cainiao relies on an intelligent logistics network to achieve the goal of 72-hour delivery of products [13]. Consumers' demands for functions such as visualization and personalization have prompted logistics companies to pay more attention to logistics service innovation.

Consumers' demands for functions such as visualization and personalization have prompted logistics companies to pay more attention to logistics service innovation. Through the intelligent logistics model, it is helpful to renew the supply chain, realize the high added value of products, high operation efficiency, and shorten the supply cycle, to obtain the active advantage of the market and realize the growth of profits [14, 15]. The intelligent transformation of logistics services is crucial to the development of an enterprise and even determines whether an enterprise can exist [3]. LSPs are usually initiated in the process of intelligent transformation of LSSCs, mergers and acquisitions guide logistics service integrators to participate through a series of incentive measures. The intelligent transformation process of LSSC is usually initiated by LSIs, and guides LSPs to participate through a series of incentive measures [16]. The intelligent transformation of logistics services can upgrade traditional logistics services. In the process, LSIs can reduce the cost of logistics services and improve operational efficiency [17], while LSPs can obtain more orders. For example, Cainiao, as one LSI of Taobao and Tmall supermarkets, integrates many functional logistics providers to provide services. Among them, YTO Express uses advanced technology and equipment in the service process, improves service processes, provides comprehensive warehousing and distribution and logistics network Expasion and other value-added ser-vices to meet product delivery. Therefore, Cainiao allocates more logistics service orders to YTO Express. However, it is a long-term task for logistics service providers to carry out intelligent logistics transformation. During the transformation process, additional values such as logistics service design, introduction of advanced technology and equipment, and maintenance all require additional costs. Uncertainty of investment and profit will be enthusiastic about the intelligent transformation of logistics service providers. Whether LSIs allow LSPs to independently trans-form intelligently or encourage LSPs to jointly intelligently transform? At the same time, in the context of the Internet economy, online platforms have become a common business model. For example, Amazon, JD.com, Suning.com in the retail industry, ransfarzl.com, aisup-port.express.cainiao.com, and huolala.cn in the logistics service industry, etc. The platform model and the resale model have become the most typical and common operating models in the Internet environment [18]. Under the platform mode, LSIs provide LSPs with places to directly connect with customers, while under the resale mode, LSIs purchase LSPs' logistics resources and serve consumers. In practice, logistics service efficiency will be a key factor for decision-makers to adopt a platform model or a resale model [19]. In the process of intelligent transformation of LSPs, whether the adoption of different models affect the transformation of intelligent logistics services?

Motivated by the above issues, this paper constructs a platform-involved LSSC, including one LSI and one LSP, we discuss the selection of intelligent transformation of the LSP under the background of the LSI launching joint logistics service intelligent transformation. This paper mainly answers the following two questions:

(1) Is the LSP going to undergo intelligent transformation? If the LSP chooses intelligently transformed, does she stand alone or join the joint intelligent transformation plan of the LSI?
(2) Are there differences in the decision-making process between the platform model and the resale model, and which model is more beneficial for both parties to choose in different decision-making processes?

The main contributions of this paper are the following three points. First, from the perspective of enterprise strategic transformation, this study not only considers consumer promotion but also considers the intelligent transformation of LSPs driven by LSIs, and analyzes the impact of intelligence level and supply chain incentives on strategic choices. Second, in this study, we have drawn many interesting conclusions about the platform model introduced under the Internet economy in the process of analyzing the intelligent transformation of logistics services. Third, this study has strong practical guiding significance for the intelligent transformation of convection services. For example, LSIs can adjust the LSPs' behavior of intelligent logistics transformation by manipulating the share ratio.

The rest of this paper is organized as follows. Section 2 reviews the relevant literature on intelligent logistics, selling mode choice and, LSSC. In section 3, under the platform mode and reselling mode, we constructed and solved the model in six scenarios including basic services provided by the LSP, intelligent transformation of independent, and joint transformation of logistics services. In section 4, we make a comparative analysis combined with numerical simulation, and section 5 concludes this paper. Finally, we give the relevant proof in Appendix A.

2 Literature review

The literature related to this paper includes the following three streams: intelligent logistics, sales mode choice and, LSSC.

The first stream lies in intelligent logistics. Intelligent logistics is a form of logistics service that ensures the entire supply chain is more intelligent and automated with the help of various advanced technologies [9]. Some scholars began to pay attention to the technology of improving the efficiency of logistics systems, such as RFID technology [20], IoT technology [21], blockchain technology [22, 23], and big data technology. The use of intelligent logistics can improve service efficiency, reduce costs, and gain first-mover advantages in the market, which is an inevitable trend in the development of modern logistics [20, 24]. Some scholars have conducted related research from the perspective of the government's policy support for the development of intelligent logistics [9, 25] and the future development trend of intelligent logistics [8, 10]. At present, the research on intelligent logistics has been extended to various scenarios. For example, the least squares method for the shortest path problem of port intelligent logistics based on cloud technology [26]; improved the algorithm problem for optimizing end-of-line delivery vehicle routing [27]; a hybrid agent scheduling and synchronization approach to solve the optimization problem of the intelligent logistics system [28], and an intelligent warehouse management approach based on machine vision and the Internet of Things (IoT) [29]. At the same time, some scholars have studied the intelligent transformation of logistics services through game theory methods [30-33]. Different from the above research, the focus of this paper is to study whether this LSP is driven by consumers to undergo intelligent transformation, or whether it is driven by the LSI alone or jointly with intelligent transformation, and mainly discusses consumption stimulation and the LSI's incentive effects.

The second stream lies in sales mode choice. Due to the development of the e-commerce economy, the comparative study on the coexistence of the platform model and the resale model has attracted more and more scholars' attention. Abhishek *et al.* [18] conducted a comprehensive study of the platform model and the resale model. Some scholars have also conducted related research on pricing and channel entry under the two models. For example, Yan *et al.* [34] considered the pricing decision of manufacturers under the platform model and the resale model and found that whether a retailer joins the platform depends on the platform fee. Zhang *et al.* [35] studied the supplier channel expansion problem in two modes. He *et al.* [36] studied the sales model and pricing of the tourism O2O supply chain. Xu *et al*. [37] researched that demand is influenced by consumers' green preferences and manufacturers' pricing and carbon emission reduction decisions under the platform model and resale model. Liu *et al*. [38] studied the different sales models and prices of fresh food, and considered channel competition and the application of blockchain technology in the process of research. Liu *et al*. [39] considered both the service model and the sales model and analyzed which model is optimal. Geng *et al*. [40] discussed the relationship between different sales models and supply chain members. Some scholars have also studied the sales system of online sales platforms [41].

In the process of product sales, logistics service is an important factor, some scholars have considered logistics service strategies in the comparative study of the two models [42-44]. The platform model and the resale model are two commonly used sales models in the e-commerce environment, but so far there are few studies on the differences in the sales model of the LSSC. This research explores the influence of logistics service intelligence level on sales mode, the influence of different modes on service cooperation between the two parties, and the relationship between share ratio and intelligent logistics service level.

The third stream lies in LSSC. Scholars' research on LSSC mainly focuses on logistics service quality, pricing, and procurement decision-making, and will consider the fair preference of supply chain members, risk aversion or social responsibility, and other behavioral factors, as well as demand uncertainty, demand update, and demand interruption, etc. factor into the model. For example, Liu *et al*. [45] studied how to determine the revenue-sharing coefficient of LSSC under random conditions. Yunmiao *et al*. [46] analyzed the problem of contract coordination selection when the demand of LSSC is uncertain. Liu *et al*. [3] studied the LSSC scheduling problem under the environment of mass customization and uncertain operation time. Liu *et al*. [47] analyzed the coordination problem of LSSC when demand is disturbed. Wang *et al*. [48] and Liu *et al*. [49] took the fairness preference factor into account in the model, and analyzed the contract coordination and order allocation decision-making problems in the LSSC, respectively. Liu *et al*. [6] discussed the optimal decision-making of member enterprises in the LSSC based on different decision-making modes under the background of "One Belt, One Road". Qin *et al*. [50] studied the service quality coordination problem of the online shopping service supply chain based on fairness and individual rational preference. Niu *et al*. [51] explored the role of the Internet of Things in the context of sustainable and traceable functions of logistics services with the help of game theory. Different from them, this paper mainly studies whether the logistics service is intelligently transformed, how to transform it, and how to sell it. Kin and Ha [52] analyzed the differences between the manufacturing supply chain and the LSSC.

According to our review of relevant literature, we found that intelligent logistics has aroused the research interest of many scholars. This literature has a great reference for us, but there are still gaps. Relevant to our study are [30, 32], they mainly focused on logistics services integration transformation contract, while we considered the transformational impact of the LSI providing cost sharing on the LSP. At the same time, they did not consider the choice of sales model during the transformation of the LSP. Combined with the above analysis, the existing research cannot solve the problem we raised. In addition, this paper also considers the intelligent transformation of logistics services and the decision-making of sales model selection and provides more theoretical decision-making references for the intelligent transformation of LSSCs.

3. Model construction and analysis

3.1 Model construction

This paper considers the LSSC involved in the platform, and its logistics service requirements are completed by one LSI and one LSP. The LSI joins the supply chain service platform and pays the franchise fee, integrates and publishes the demand information of logistics services, and the LSP completes the tasks according to the needs of customers. The modes adopted when the LSI and the LSP cooperate include resale mode and platform mode. In the platform mode, the LSP provides services to customers in the form of third-party logistics according to the requirements of the LSI and pays u proportion of the commission, as shown in Fig. 1. In the resale mode, the logistics service is wholesaled to the LSI at the wholesale price w, and the LSI determines the final logistics service price p, as shown in Fig. 2. In the above two modes, the LSP will consider whether to carry out intelligent transformation according to the requirements of the LSI. In the case of independent transformation, the LSP shall bear the cost of intelligent transformation, and in the case of joint transformation, the LSI shall share s proportion of the transformation cost.

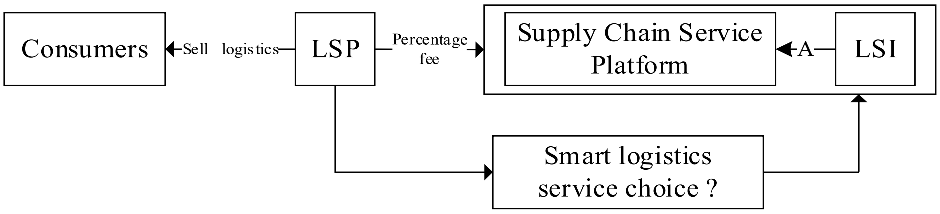


Fig. 1 Platform mode

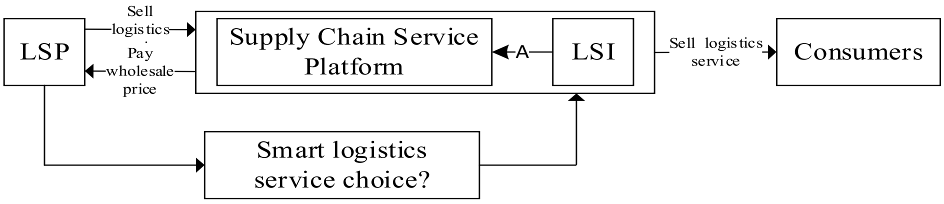


Fig. 2 Resale mode

According to the literature [30, 31], we assume that the cost of intelligent transformation of logistics services is $c(y) = gy^2$, which reflects that with the increase of intelligent logistics service level, the cost will also increase. Among them, g represents the cost-sensitive factor of intelligent logistics service, and the larger the value, the more sensitive it is. We denote $z = m^2/g$ as the efficiency of intelligent logistics service, reflecting the ratio of benefit to cost per unit of intelligent logistics service efficiency. Without loss of generality, similar to the literature [30], we also assume that the marginal cost of logistics services is 0, that is, this paper does not consider the production cost of unit logistics services and other costs other than intelligent logistics services. In the platform model, the revenue share of the LSI is related to the category of logistics services. This paper assumes that the revenue share ratio u is an exogenous variable. To conform to the reality and the profit of the LSI is positive, when the LSI is not provided and the LSP undertakes intelligent logistics transformation alone, the value range of u is $0 < u < 0.5$. When implementing joint intelligent transformation, to ensure that the cost shared by logistics service integrators is positive, the following conditions need to be met: $\frac{-6-2\sqrt{9-2z}+z}{z} < u < 0.5$ and $0 < z < 2$.

Demand. According to the literature [31, 33], when the customer demand is the basic logistics service, the demand function can be expressed as $q = a - p$, among them, a is the basic demand of logistics services, and p is the price of logistics services. In fact, in the actual logistics service operation process, in addition to the basic logistics needs, customers will also require personalized value-added logistics services. To meet the needs of customers, the LSI will propose intelligent logistics transformation, and the LSP will specifically undertake intelligent logistics service tasks. Assuming that under the influence of intelligent logistics, the increase in logistics demand at this time is my, among them, y represents the service level of intelligent logistics, and m represents the sensitivity factor of customers to intelligent logistics services, Therefore, the demand for intelligent logistics services can be expressed as $q = a - p + my$.

Scenarios. Considering the strategic choice of the LSP between the platform model and the resale model, as well as the choice of the intelligent transformation strategy in the face of customer needs, there are 6 possible decision scenarios, as shown in Table 1.

Table 1 Logistics service and sales choice mode in six scenarios

Sales mode	Marketing type		
	Basic logistics	Intelligent logistics by the LSP	Joint Intelligent logistics
Platform mode	PB	PS	PC
Resale mode	RB	RS	RC

- Platform mode with basic logistic service (PB): In this case, the LSI acts as a bridge between consumers and the LSP, and the LSP directly provides consumers with basic logistics services and decides the price p^{PB} of logistics services.
- Platform mode with intelligent logistic service by the LSP (PS): In this case, the LSI proposes a joint intelligent logistics service transformation plan, and the LSP does not participate in the joint plan and decides the transformation level y^{PS} and sales price p^{PS} independently.
- Platform mode with intelligent logistic service by cooperation (PC): In this case, the LSI proposed a joint intelligent logistics service transformation plan, and the logistics service provider decided to participate. First, the LSI decides to share the cost of intelligent logistics transformation, and then the LSP decides the level of intelligent logistics service transformation y^{PC} and sales price p^{PC}.
- Resale mode with basic logistic service (RB): In this case, the LSP resells the service to the LSI. First, the LSP decides the basic logistics service wholesale price w^{RB}, and the LSI decides the final logistics service price p^{RB}.
- Resale mode with intelligent logistic service by LSI (RS): In this case, the LSI proposes a joint intelligent logistics service transformation plan, and the LSP does not participate. First, the LSP decides the wholesale price w^{RS}, then the LSI decides the sales price p^{RS}, finally the LSP decides the intelligence level y^{RS}.
- Resale mode with intelligent logistic service by cooperation (RC): In this case, the LSI proposed a joint intelligent logistics service transformation plan, and the LSP decided to participate. First, the LSP decides the wholesale price w^{RC}, then the LSI decides the sales price p^{RC} and cost-sharing ratio s^{RC}, and finally the LSP decides the intelligence level y^{RC}.

The main notations related to this paper are shown in Table 2. We use the superscript P and R to represent the platform sales model and the resale model, respectively, the superscript *Z to represent the optimal solution, and π_T and π_J to represent the profits of the LSP and the LSI, respectively.

Table 2 The notations related to this paper

Notation	Instruction
a	Market size of the logistics service market
p	Price of the logistics service
q	Demand for the logistics service
m	Consumer's sensitivity of the intelligent logistics service
y	Intelligent level of the logistics service
g	Cost coefficient of intelligent transformation
u	The share ratio of the LSI under a platform mode
w	Wholesale price of the unit logistics service under a resale mode
z	The intelligent transformation efficiency
s	Cost sharing ratio of the intelligent transformation
A	Supply chain platform membership fee
$\pi_T^{PB}(\pi_T^{RB})$	Profit of the LSP in platform mode (resale mode) with basic service
$\pi_J^{PB}(\pi_J^{RB})$	Profit of the LSI in platform mode (resale mode) with basic service
$\pi_T^{PS}(\pi_T^{RS})$	Profit of the LSP in platform mode (resale mode) with intelligent transformation by the LSP
$\pi_J^{PS}(\pi_J^{RS})$	Profit of the LSI in platform mode (resale mode) with intelligent transformation by the LSP
$\pi_T^{PC}(\pi_T^{RC})$	Profit of the LSP in platform mode (resale mode) with intelligent transformation by cooperation
$\pi_J^{PC}(\pi_J^{RC})$	Profit of the LSI in platform mode (resale mode) with intelligent transformation by cooperation

3.2 Equilibrium analysis

In this section, we solve and analyze the equilibrium results in six scenarios, where the solution is through the reverse recursion method. To ensure that the members of the LSSC service supply chain have positive benefits in each case, we assume that $4 - z > 0$.

PB scenario

In the PB scenario, the LSI serves as a connector between the LSP and consumers, and the LSP provides basic logistics services. At this point, the profit functions of the LSP and the LSI are:

$$\begin{cases} max\, \pi_T^{PB} = (1-u)p(a-p) \\ max\, \pi_J^{PB} = up(a-p) - A \end{cases} \tag{1}$$

In this scenario, the LSP first determines the basic logistics service price to maximize its profit, and then she pays a certain percentage of transaction fees to the LSI. We can get the following equilibrium results in the PB scenario.

$$p^{PB*} = \frac{a}{2}, q^{PB*} = \frac{a}{2}, \pi_T^{PB*} = \frac{a^2(1-u)}{4}, \pi_J^{PB*} = \frac{a^2 u}{4} - A.$$

It is easy to see from the equilibrium solution that the optimal logistics service price, the logistics service market demand, the LSI's and the LSP's profit increase with the increase of the basic logistics service market. At the same time, the profits of the LSP and the LSI decrease and increase respectively with the increase of the share ratio.

PS scenario

In the PS scenario, although the LSI proposed a joint logistics service intelligent transformation plan, but the LSP did not join the plan, at this time, the logistics service intelligent transformation cost is borne by the LSP. At this point, the profit functions of the LSP and the LSI are:

$$\begin{cases} max\, \pi_T^{PS} = (1-u)p(a-p+my) - gy^2 \\ max\, \pi_J^{PS} = up(a-p+my) - A \end{cases} \tag{2}$$

In this scenario, in addition to determining the price of logistics services, the LSP also needs to decide the intelligence level of logistics service. We can get the following equilibrium results in the PS scenario.

$$p^{PS*} = \frac{2a}{4-z(1-u)}, q^{PS*} = \frac{am(1-u)}{4g+m^2u-m^2}, y^{PS*} = \frac{2a}{4-z(1-u)}, \pi_T^{PS*} = \frac{a^2(1-u)}{4-z(1-u)}, \pi_J^{PS*} = \frac{4a^2u}{[4-z(1-u)]^2} - A.$$

Obviously, with the increase in the demand for basic logistics services and the efficiency of intelligent logistics transformation, the price, the market demand, the intelligence level of logistics services, and the profits of the LSI and the LSP have all increased. With the increase of the share ratio, the price, the market demand, and the intelligence level of logistics services, the profit of the LSP decreases and the profit of the LSI increases.

PC scenario

In the PC scenario, the LSP joins the joint logistics service intelligent transformation plan proposed by the LSI, and the LSI bears the proportion of the cost of the intelligent transformation as s. At this point, the profit functions of the LSP and the LSI are:

$$\begin{cases} max\, \pi_T^{PC} = (1-u)p(a-p+my) - (1-s)gy^2 \\ max\, \pi_J^{PC} = up(a-p+my) - sgy^2 - A \end{cases} \tag{3}$$

In this scenario, the LSP first determines the logistics service price p and the logistics service intelligence level y, and then the LSI decides the final cost-sharing ratio s of intelligent logistics transformation. We can get the following equilibrium results in the PC scenario.

$$p^{PC*} = \frac{a[8-z(1-u)]}{4(4-z)}, q^{PC*} = \frac{a[8-z(1-u)]}{4(4-z)}, \ y^{PC*} = \frac{am(u+1)}{8g-2m^2},$$

$$s^{PC*} = \frac{z(1-u)^2 + 4(3u-1)}{4(1+u)}, \pi_T^{PC*} = \frac{a^2(1-u)[8-z(1-u)]}{8(4-z)}, \pi_J^{PC*} = \frac{a^2[z(1-u)^2 + 16u]}{16(4-z)} - A.$$

Similar to the PB scenario, with the increase in the demand for basic logistics services and the efficiency of intelligent transformation, the price, the market demand, the intelligence level of logistics services, the profits of the LSI and the LSP all increase. With the increase of the share ratio, the price, the market demand, the intelligence level of logistics services, and the profits of the LSP all decrease while the profits of the LSI increase.

RB scenario

In the RB scenario, the LSI also acts as an intermediary, and the basic logistics service is still performed by the LSP. At this point, the profit functions of the LSP and the LSI are:

$$\begin{cases} max\, \pi_T^{RB} = w(a-p) \\ max\, \pi_J^{RB} = (p-w)(a-p) - A \end{cases} \tag{4}$$

In this scenario, first, the LSP decides the wholesale price, and then the LSI decides the sales price. We can get the following equilibrium results in the RB scenario.

$$p^{RB*} = \frac{a}{2}, q^{RB*} = \frac{3a}{4}, w^{RB*} = \frac{a}{4}, \pi_T^{RB*} = \frac{a^2}{8}, \pi_J^{RB*} = \frac{a^2}{16} - A.$$

It is easy to see that with the increase in the market demand for basic logistics services, the wholesale price, the sales price, the market demand, and the profits of the LSP and the LSI all increase.

RS scenario

In the RS scenario, the LSP does not participate in the joint logistics service intelligent transformation plan proposed by the LSI, and the LSI's share of the cost of the logistics service intelligent transformation is 0. At this point, the profit functions of the LSP and the LSI are:

$$\begin{cases} max\, \pi_T^{RS} = w(a-p+my) - gy^2 \\ max\, \pi_J^{TS} = (p-w)(a-p+my) - A \end{cases} \tag{5}$$

In this scenario, the LSP first determines the wholesale price, and then the LSP decides the sales price of the logistics service and the intelligence level. We can get the following equilibrium results in the RS scenario.

$$w^{RS*} = \frac{a}{2}, p^{RS*} = \frac{a(z+6)}{8}, q^{RS*} = \frac{a(2+z)}{8}, y^{RS*} = \frac{am}{4g}, \pi_T^{RS*} = \frac{a^2}{8}, \pi_J^{RS*} = \frac{a^2(z+2)^2}{64} - A.$$

Obviously, with the increase in the demand for basic logistics services, the wholesale price, the sales price, the market demand, the level of intelligence of logistics services, the profits of the LSI and the LSP have all increased. With the increase of the efficiency level of intelligent logistics service transformation, the sales price, the market demand, and the profits of the LSI increase. The intelligence level of logistics services is positively correlated with consumers' preferences and negatively correlated with the cost coefficient of intelligent transformation.

RC scenario

In the RC scenario, the LSP participates in the intelligent logistics joint transformation plan proposed by the LSI, and the LSI shares s proportion of the cost of intelligent transformation. At this point, the profit functions of the LSP and the LSI are:

$$\begin{cases} max\, \pi_T^{RC} = w(a-p+my) - (1-s)gy^2 \\ max\, \pi_J^{RC} = (p-w)(a-p+my) - sgy^2 - A \end{cases} \tag{6}$$

In this scenario, first, the LSP decides the wholesale price, then the LSP decides the sales price and cost allocation ratio of the logistics service, and finally, the LSP decides the intelligence level. We can get the following equilibrium results in the RC scenario.

$$w^{RC*}=\frac{a}{2},p^{RC*}=\frac{a(12-z)}{4(4-z)},q^{RC*}=\frac{a(z+4)}{4(4-z)},y^{RC*}=\frac{am}{4g-m^2},s=\frac{z}{4},\pi_T^{RC*}=\frac{a^2}{8},\pi_J^{RC*}=\frac{a^2(3z+4)}{16(4-z)}-A.$$

It is easy to see that the basic demand for logistics services and the efficiency of intelligent logistics service transformation have a similar impact on the equilibrium results in this scenario as in the RS scenario. At the same time, the influence of consumers' intelligent logistics preference and logistics service intelligent transformation cost coefficient on the intelligence level of logistics service is consistent with the impact of the RS scenario, and the proportion of intelligent logistics transformation cost sharing is positively related to the efficiency of intelligent logistics service transformation.

In the next section, we first analyze the intelligent logistics strategy choices under the platform mode and the resale mode and then compare the differences in intelligent logistics decision-making under the two modes to provide relevant management insights and implications.

4. Results and insights

4.1 Analysis under the platform mode

By comparing the equilibrium logistics service sales price, logistics service market demand, logistics service intelligence level, and profit of the LSP and the LSI under the three scenarios of the platform model, we can get Corollary 1.

Corollary 1: (1) $p^{PB*}<p^{PS*}<p^{PC*}, q^{PB*}<q^{PS*}<q^{PC*}, y^{PS*}<y^{PC*}$. (2) $\pi_T^{PB*}<\pi_T^{PS*}<\pi_T^{PC*}$ if $0<z<1$ and $\pi_T^{PB*}<\pi_T^{PC*}<\pi_T^{PS*}$ if $1<z<2$; $\pi_J^{PB*}<\pi_J^{PS*}<\pi_J^{PC*}$.

To intuitively see the size of the equilibrium results in the three cases, we assume $a=1$. Combined with the numerical simulation results in Figs. 3-7, it can be seen that the comparison of the equilibrium results in the three cases under the platform mode is consistent with Corollary 1.

Corollary 1(1) shows that under the platform model, the implementation of intelligent logistics transformation will increase the price and sales of logistics services and the logistics service intelligence level is the highest during the joint logistics service intelligence transformation plan. The equilibrium price of the joint intelligent logistics service is the highest during the transformation. The reason for this phenomenon is that the LSP invests more in improving the intelligent level of logistics services, and the LSP transfers to consumers through price increases, make up for the extra costs incurred. The demand for the intelligent transformation of joint logistics services is also the highest. At this time, the demand for logistics services is simultaneously affected by the price and the intelligence level of logistics services. The intelligent transformation of joint logistics services brings a more obvious demand increase.

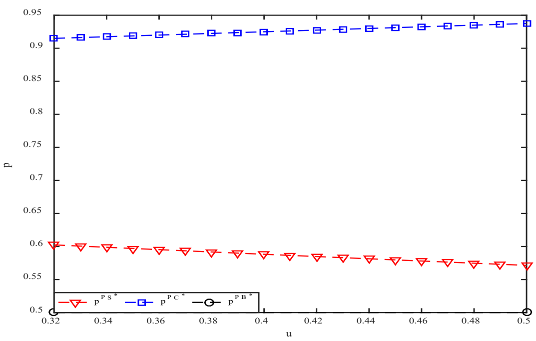


Fig. 3 Comparison of p under the platform mode

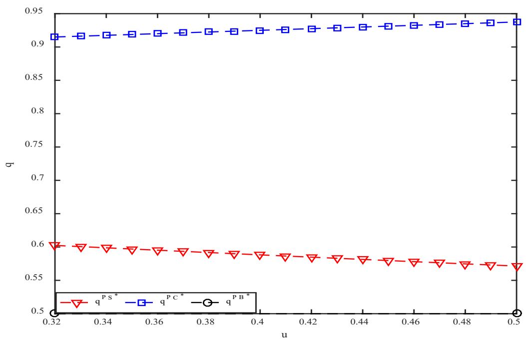


Fig. 4 Comparison of q under the platform mode

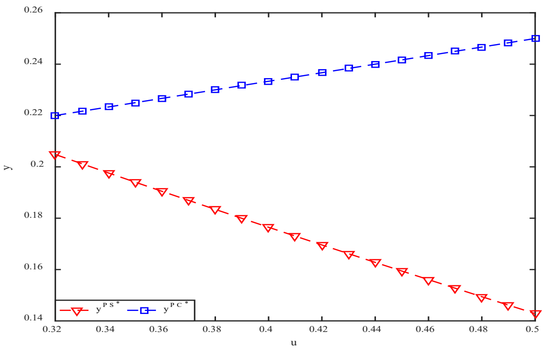


Fig. 5 Comparison of y under the platform mode

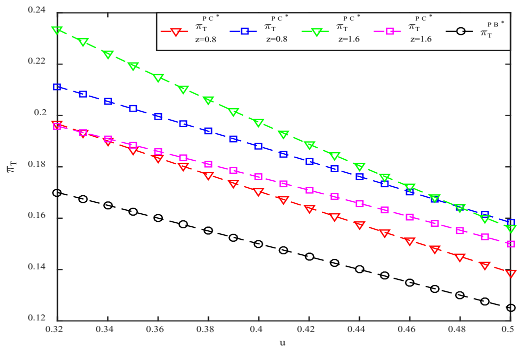


Fig. 6 Comparison of π_T under the platform mode

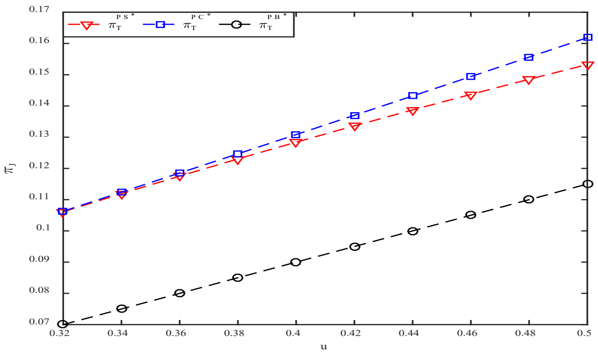


Fig. 7 Comparison of π_J under the platform mode

From Corollary 1(2), we know that the LSP participating in the LSIs' intelligent transformation will get more profit. The efficiency of logistics service intelligence transformation affects the benefit of the LSP when she adopts individual or joint logistics service intelligence. That is, when the efficiency of intelligent transformation is low, it is beneficial to the LSP when it is jointly transformed, and vice versa, it is beneficial to the LSP when it is transformed alone. For the LSI, she will get more benefits when carrying out joint intelligent transformation. If the efficiency of intelligent transformation is low, the benefit brought by the LSP's independent logistics service intelligent transformation will not be significant, and the plan to join the LSI's plan will get a part of the cost-sharing, so that the LSP's profit is higher than that of the single intelligent logistics transformation. If the efficiency of logistics service intelligence is high, although the joint transformation plan can improve the intelligence level and share a part of the cost, because of the higher share cost, the LSP is more willing to choose a separate intelligent transformation. For the LSI, choosing the joint intelligent transformation is the optimal decision. Because the LSI can grasp the different needs of consumers, and encourage the LSP to join the plan according to consumers' needs, to provide more satisfactory logistics services.

4.2 Analysis under the resale mode

Similar to the platform model, in this section, we compare the equilibrium wholesale price, the sale price, the market demand, the intelligence level, and the profit of the logistics service supply chain under the three scenarios of the resale model, we can get Corollary 2.

Corollary 2: (1) $w^{RB*} = w^{RS*} = w^{RC*}$, $p^{RB*} < p^{RS*} < p^{RC*}$, $y^{PS*} < y^{PC*}$, $q^{RB*} < q^{RS*} < q^{RC*}$.
(2) $\pi_T^{RB*} = \pi_T^{RS*} = \pi_T^{RC*}, \pi_J^{RB*} < \pi_J^{RS*} < \pi_J^{RC*}$.
Combined with the numerical simulation results in Figs. 8-11, it can be seen that the comparison relationship between the equilibrium results in the three scenarios is consistent with Corollary 2. The relationship between the wholesale price and the profit of the LSP in the three scenarios can be seen from the comparison in Table 3, and the numerical simulation is omitted.

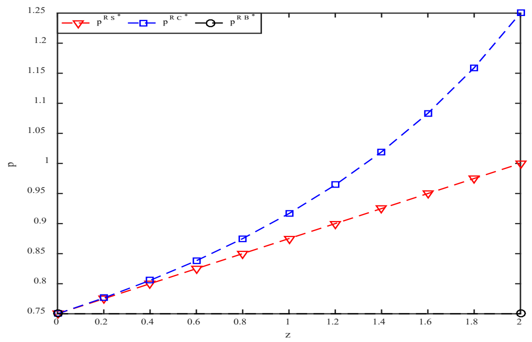


Fig. 8 Comparison of p under the resale mode

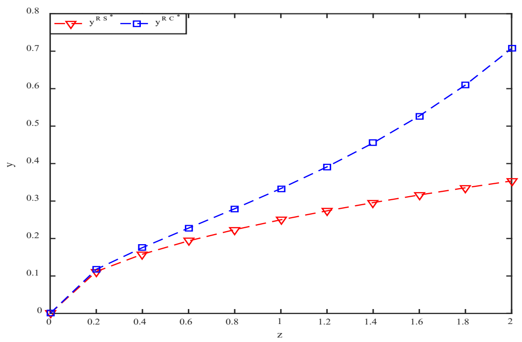


Fig. 9 Comparison of y under the resale mode

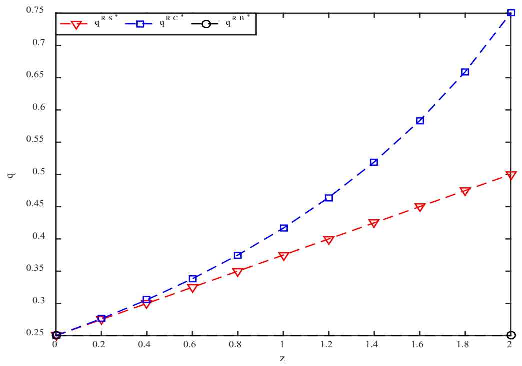


Fig. 10 Comparison of q under the platform mode

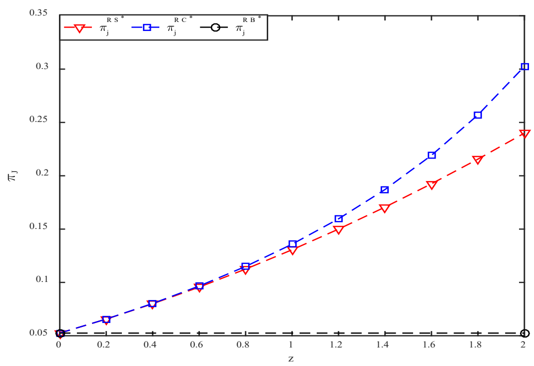


Fig. 11 Comparison of π_J under the platform mode

Corollary 2 shows that the equilibrium logistics service price, the level of logistics service intelligence, the market demand, and the LSI's profit are optimal when the LSP participates in the intelligent transformation under the resale mode, which is consistent with the platform mode. Under the resale model, the wholesale price and the LSP's profit in the three scenarios remain unchanged, which is different from the conclusion under the platform model. The main reason may be that the intelligent transformation in the resale model brings more demand but also costs more, and the cost at this time is offset by the increase in revenue brought by the increase in demand.

Therefore, the joint intelligent transformation plan of the LSI under the resale model is not very attractive to the LSP. In the resale mode, the choice of the LSP to choose independent or joint logistics service intelligent transformation depends on the efficiency of logistics service intelligent transformation. That is to say, the LSI will gain more benefits during joint intelligence transformation, that is, the LSI should actively encourage the LSP to participate in the joint intelligent transformation plan.

4.3 Comparative analysis

In this section, we will compare and analyze the equilibrium results of the LSP under different service strategies under the platform model and the resale model. We can obtain Corollary 3, Corollary 4, and Corollary 5.

Corollary 3. Comparing the equilibrium solutions of the two modes when the LSP provides basic logistics services, we can get

(1) $p^{PB*} < p^{RB*}, q^{PB*} > q^{RB*}$. (2) $\pi_T^{PB*} < \pi_T^{RB*}$ if $0 < u < \frac{1}{2}$ and $\pi_T^{PB*} > \pi_T^{RB*}\ \frac{1}{2} < u < 1$; $\pi_J^{PB*} < \pi_J^{RB*}$ if $0 < u < \frac{1}{4}$ and $\pi_J^{PB*} > \pi_J^{RB*}$ if $\frac{1}{4} < u < 1$.

According to the equilibrium results in Table 3, it is easy to conclude in Corollary 3, so we omit the numerical simulation results. According to Corollary 3(1), when the LSP provides basic logistics services, the platform mode has lower logistics service prices and higher logistics service demand. The reason is that under the platform model, the price is directly determined by the LSP, which avoids the price increase by layers of pricing. According to Corollary 3(2), when the share ratio is lower than 0.5, the LSP is more willing to adopt the platform model, and when the share ratio is less than 0.25, the LSI is more willing to adopt the resale model. At the same time, when the split ratio is between 0.25 and 0.5, the platform mode selection is the optimal strategy for both. That is to say, when providing basic logistics services, the choice of the sales model for both parties mainly depends on the size of the share ratio.

Corollary 4. Comparing the equilibrium solutions of the two modes when the logistics service provider independently carries out the intelligent transformation, we can obtain:

(1) $p^{PS*} < p^{RS*}$, $q^{PS*} > q^{RS*}$; If $0 < u < \frac{z}{4+z}$, $y^{PS*} > y^{RS*}$ and $y^{PS*} < y^{RS*}$ if $\frac{z}{4+z} < u < \frac{1}{2}$.
(2) $\pi_T^{PS*} > \pi_T^{RS*}$. When $0 < u < \frac{1}{4}$, $\pi_J^{PS*} < \pi_J^{RS*}$; and when $\frac{1}{4} < u < \frac{1}{2}$, $\pi_J^{PS*} > \pi_J^{RS*}$ if $0 < z < z_0$; $\pi_J^{PS*} < \pi_J^{RS*}$ if $z_0 < z < 2$, where $z_0 = -\frac{2(9+6u+u^2)}{3(-7+6u+u^2)}$.

Combined with the numerical simulation results in Figs. 12-16, it can be seen that the equilibrium solutions in the two modes are consistent with Corollary 4 when the logistics service provider independently transforms its logistics service intelligence.

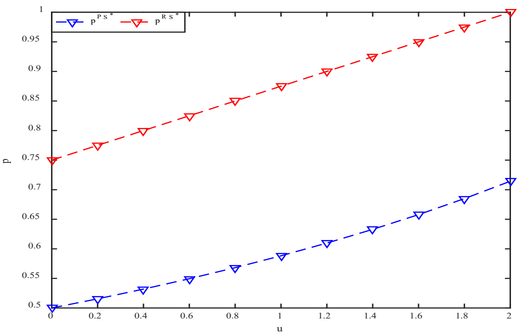


Fig. 12 Comparison of p under between PS and RS scenarios

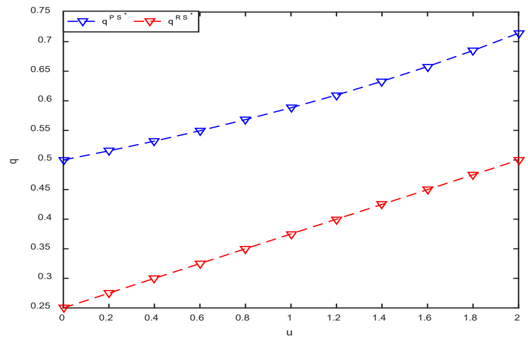


Fig. 13 Comparison of $q$between PS and RS scenarios

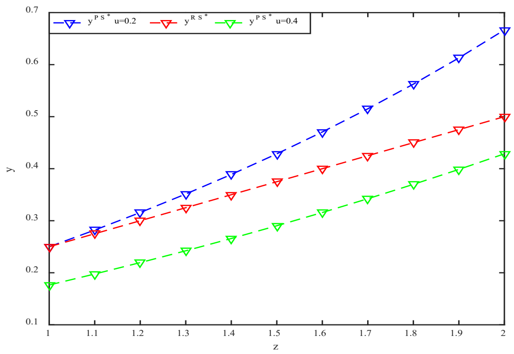


Fig.14 Comparison of y under between PS and RS scenarios

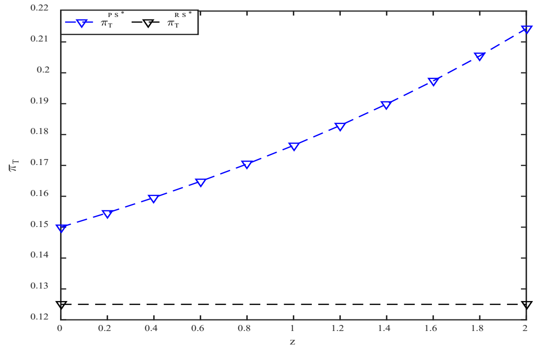


Fig. 15 Comparison of π_T between PS and RS scenarios

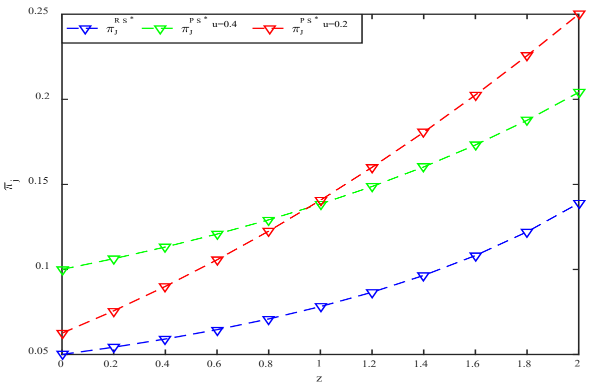


Fig. 16 Comparison of π_J between PS and RS scenarios

Corollary 4(1) shows that the logistics service price is lower in the platform mode when the LSP is independently intelligently transformed. The main reason is that there is no layer-by-layer price increase phenomenon, and the LSP will adopt measures of small profits but quick turnover to obtain more benefits. The intelligence level of logistics services under the two modes is related to the size of the share ratio. When the share ratio charged by the LSI is low, the LSP has the motivation to carry out the intelligent transformation. On the contrary, the enthusiasm of the LSP for intelligent transformation is not high, so the intelligent level of the logistics service in the platform mode is low.

Corollary 4(2) shows that the LSP can obtain more benefits by adopting the platform model when she chooses independently intelligent transformation. Although the LSP needs to pay a certain percentage of the share under the platform model, the market demand for logistics services has decreased within a certain range, but it is still higher than the resale model, and the double marginal problem caused by the increase in prices is avoided. The profit of the LSI under the two modes is related to the share ratio and the efficiency of intelligent transformation. When the share ratio is low, its income is greatly reduced, at this time, the LSI tends to choose the resale model. When the share ratio is high, if the efficiency of intelligent transformation is low, it is more advantageous for the LSI to choose the platform mode, and the platform mode is the optimal mode for both parties. If the efficiency of intelligent transformation is low, the LSI tends to choose the resale model.

Therefore, the determination of the final sales model and share ratio is affected by the market position and business negotiation strategies of both parties. At the same time, the logistics service intelligence level needs to be paid attention to. Because both parties conduct a series of commercial and technical activities to meet consumer preferences. By evaluating consumers' satisfaction with the effect of intelligent transformation, it can effectively coordinate the differences in goals between the two parties and make more scientific decisions in advance between the benefits of intelligent transformation and the losses caused by business model transformation.

Corollary 5. Comparing the equilibrium solutions of the two modes when the LSP chooses joint intelligent transformation, we can get:

(1) $y^{PC*} < y^{RC*}$, $p^{PC*} < p^{RC*}$,$q^{PC*} > q^{RC*}$. (2) $\pi_T^{PC*} > \pi_T^{RC*}$;When $\frac{1}{4} < u < \frac{1}{3}$, if $\frac{4(1-4u)}{-2-2u+u^2} < z < \frac{4(1-3u)}{1-2u+u^2}$, $\pi_J^{PC*} > \pi_J^{RC*}$ and if $\max\{\frac{4(1-3u)}{1-2u+u^2}, \frac{4(1-4u)}{-2-2u+u^2}\} < z < 2$,$\pi_J^{PC*} < \pi_J^{RC*}$; When $\frac{1}{3} < u < \frac{1}{2}$, if $0 < z < \frac{4(1-4u)}{-2-2u+u^2}$, $\pi_J^{PC*} > \pi_J^{RC*}$; and if $\frac{4(1-4u)}{-2-2u+u^2} < z < 2$, $\pi_J^{PC*} < \pi_J^{RC*}$.

Combined with the numerical simulation results in Figs. 17-21, it can be seen that the equilibrium solution results in the two modes are consistent with Corollary 5 when the LSP chooses joint intelligent transformation.

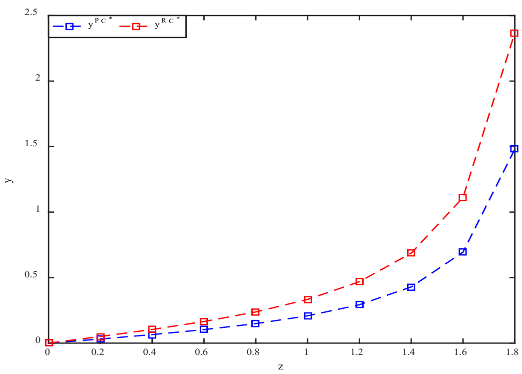


Fig. 17 Comparison of y under between PC and RC scenarios

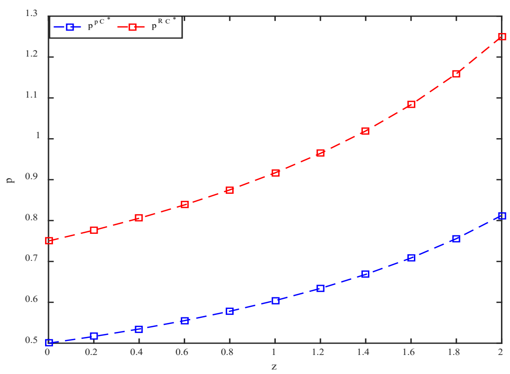


Fig. 18 Comparison of p between PC and RC scenarios

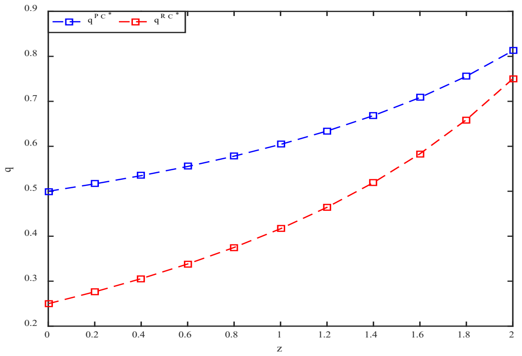


Fig. 19 Comparison of q under between PC and RC scenarios

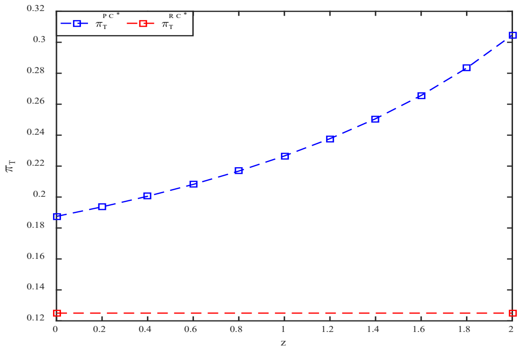


Fig. 20 Comparison of $\pi_T$between PC and RC scenarios

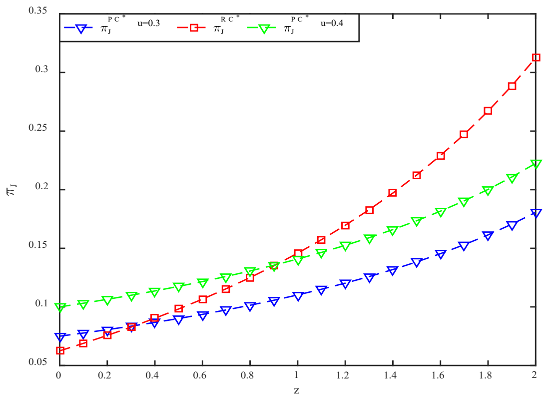


Fig. 21 Comparison of π_J under between PC and RC scenarios

Corollary 5(1) shows that when the LSP chooses joint intelligent transformation, the logistics service intelligence level under the resale mode is higher. On the one hand, it is because the cost of intelligent transformation is shared under the resale model, and on the other hand, the LSP can obtain the benefits of intelligent transformation of intelligent logistics services. In the process of intelligent transformation, the additional cost of logistics service providers in the resale mode is transferred to the LSI through wholesale prices, so the LSI will appropriately increase the sale price.

Corollary 5(2) shows that when the LSP chooses to participate in joint intelligent transformation, she will gain more benefits under the platform model. Since the double marginal problem of cooperation between the two parties in the platform model is eliminated, the resale model will make the price of logistics services too high, resulting in a decrease in the demand for the entire logistics service market. For the LSI, the profit of the LSI in the two modes depends on the proportion of shares and the efficiency of intelligent logistics service transformation. When the efficiency of intelligent transformation is low and the share ratio is not too low, the profit of the LSI in the platform mode is higher than that in the resale mode. Conversely, the LSI will gain more benefits under the resale model. Therefore, when the LSP chooses intelligent transformation, the LSI will decide which mode to adopt according to the efficiency of the logistics service transformation and the share ratio obtained.

5. Conclusion

5.1 Key findings and managerial implications

In the context of intelligent logistics transformation, this paper discussed whether the LSP should transform alone or participate in the joint transformation of the LSI's plan, and compared the difference of equilibrium results under the platform model and the resale model. We got the following conclusions.

First, when providing basic logistics services, the platform model has lower prices and a higher sales volume, and the profits of the LSP and the LSI depend on the size of the share ratio. Meanwhile, choosing the intelligent transformation can improve the benefit of the entire supply chain. Second, when the LSP chooses joint intelligent transformation, she can obtain more market demand with higher intelligent logistics service level, and at the same time, the LSI can obtain more profits. For the LSP, under the platform model, the LSP can obtain higher profit, while under the resale model, regardless of whether she participates in the joint transformation, there is no difference in her profit. Third, for the LSP, when participating in the joint intelligent transformation, the level of intelligent logistics service in the resale mode is higher than that in the platform mode; when the LSP independently transforms its logistics service intelligently, the share ratio will affect the level of intelligent logistics service under the two modes. Whether the LSP conducts independent or joint intelligent transformation, more profits can be obtained in the agency mode. Finally, for the LSI, no matter which intelligent transformation method the LSP chooses, the share ratio and the level of logistics service intelligence will affect the size of her profit under the two modes.

Based on the research in this paper, we can draw the following management insights. First, the intelligent transformation of logistics services is necessary for LSP, because more profit can be obtained with a higher level of logistics service intelligence.
Secondly, the LSP should actively participate in joint intelligent transformation. Although there is no difference in profit whether or not to participate in the joint intelligent transformation under the resale mode, the LSP can request compensation from other aspects to ensure the overall intelligence level of logistics services and its own profit. Finally, in the platform mode, the LSI should consider the intelligent transformation willingness of the LSI when adopting joint intelligent transformation. When there is a deviation in the preferences of both parties, it is necessary to balance the effects brought by the intelligent transformation and the risks brought by the choice of sales models.

5.2 Limitations and future research directions

In addition, there are still some limitations in this study, which can be used as a direction for further attention in the future. First, this paper considers cost sharing in the joint intelligent transformation, and future research is worthy of further research on other forms of compensation. Second, we assume that the information of the LSI and the LSP is completely symmetrical, and it will be necessary to consider asymmetric information for intelligent transformation. Finally, we assume that the impact of intelligent transformation on demand is certain, and the conclusions drawn from studying the impact of intelligent transformation on other forms of demand may be more interesting.

Acknowledgement

The study is supported by a project funded by Fundamental Funds for Humanities and Social Sciences Program of the Ministry of Education, (20YJC630124,21YJC630077), Technologies Research and Development Program of Henan Province (212102310998), Humanities and Social Sciences Project of Universities in Henan Province (2021-ZDJH-402), Philosophy and Social Science Planning Project of Henan Province (2021BJJ104), and Youth Backbone Cultivation Program for Colleges and Universities in Henan Province (2020GGJS175). We appreciate their support very much.

References

[1] Chandra, C., Kamrani, A. (2004). *Mass customization: A supply chain approach*, Springer Science+Business Media New York, USA, doi: 10.1007/978-1-4419-9015-0.

[2] Choy, K.L., Li, C.-L., So, S.C.K., Lau, H., Kwok, S.K., Leung, D.W.K. (2007). Managing uncertainty in logistics service supply chain, *International Journal of Risk Assessment and Management*, Vol. 7, No. 1, 19-43, doi: 10.1504/IJRAM.2007.011408.

[3] Liu, W., Wang, Q., Mao, Q., Wang, S., Zhu, D. (2015). A scheduling model of logistics service supply chain based on the mass customization service and uncertainty of FLSP's operation time, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 83, 189-215, doi: 10.1016/j.tre.2015.09.003.
[4] Liu, W., Liu, C., Ge, M. (2013). An order allocation model for the two-echelon logistics service supply chain based on cumulative prospect theory, *Journal of Purchasing and Supply Management*, Vol. 19, No. 1, 39-48, doi: 10.1016/j.pursup.2012.11.005.
[5] Liu, W.-H., Xu, X.-C., Ren, Z.-X., Peng, Y. (2011). An emergency order allocation model based on multi-provider in two-echelon logistics service supply chain, *Supply Chain Management,* Vol. 16, No. 6, 391-400, doi: 10.1108/13598541111171101.
[6] Liu, X., Zhang, K., Chen, B., Zhou, J., Miao, L. (2018). Analysis of logistics service supply chain for the One Belt and One Road initiative of China, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 117, 23-39, doi: 10.1016/j.tre.2018.01.019.
[7] Wang, J., Lim, M.K., Zhan, Y., Wang, X.-F. (2020). An intelligent logistics service system for enhancing dispatching operations in an IoT environment, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 135, Article No. 101886, doi: 10.1016/j.tre.2020.101886.
[8] Winkelhaus, S., Grosse, E.H. (2020). Logistics 4.0: A systematic review towards a new logistics system, *International Journal of Production Research,* Vol. 58, No. 1, 18-43, doi: 10.1080/00207543.2019.1612964.
[9] Pan, X., Li, M., Wang, M., Zong, T., Song, M. (2020). The effects of a smart logistics policy on carbon emissions in China: A difference-in-differences analysis, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 137, Article No. 101939, doi: 10.1016/j.tre.2020.101939.
[10] Liu, W., Liang, Y., Bao, X., Qin, J., Lim, M.K. (2022). China's logistics development trends in the post COVID-19 era, *International Journal of Logistics Research and Applications,* Vol. 25, No. 6, 965-976, doi: 10.1080/13675567.2020.1837760.
[11] Li, X. (2015). Innovation center promotes advances in logistics, from *http://www.chinadaily.com.cn/world/2015-05/20/content_20776112.htm* , accessed May 20 2015.
[12] Lin, C.Y. (2020). In the face of lockdown, China's e-commerce giants deliver, Harvard Business Review, from *https://hbr.org/2020/04/in-the-face-of-lockdown-chinas-e-commerce-giants-deliver.html*, accessed April 01, 2020.
[13] Campbell, C. (2020). China's Cainiao is revolutionizing how goods get delivered. Will the rest of the world follow its rules?, TIME, from *https://time.com/5914173/cainiao-logisitics-alibaba-china-trade/*, accessed November 23, 2020.
[14] Dan, B., Zhang, S., Zhou, M. (2018). Strategies for warranty service in a dual-channel supply chain with value-added service competition, *International Journal of Production Research*, Vol. 56, No. 17, 5677-5699, doi: 10.1080/00207543.2017.1377355.
[15] Zhong, R.Y., Huang, G.Q., Lan, S., Dai, Q.Y., Chen, X., Zhang, T. (2015). A big data approach for logistics trajectory discovery from RFID-enabled production data, *International Journal of Production Economics*, Vol. 165, 260-272, doi: 10.1016/j.ijpe.2015.02.014.
[16] Yan, X., Liu, W., Lim, M.K., Lin, Y., Wei, W. (2022). Exploring the factors to promote circular supply chain implementation in the smart logistics ecological chain, *Industrial Marketing Management*, Vol. 101, 57-70, doi: 10.1016/j.indmarman.2021.11.015.
[17] Fan, T., Pan, Q., Pan, F., Zhou, W., Chen, J. (2020). Intelligent logistics integration of internal and external transportation with separation mode, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 133, Article No. 101806, doi: 10.1016/j.tre.2019.10.011.
[18] Abhishek, V., Jerath, K., Zhang, Z.J. (2015). Agency selling or reselling? Channel structures in electronic retailing, *Management Science*, Vol. 62, No. 8, 2259-2280, doi: 10.1287/mnsc.2015.2230.
[19] Tian, L., Vakharia, A.J., Tan, Y., Xu Y. (2018). Marketplace, reseller, or hybrid: Strategic analysis of an emerging e-commerce model, *Production and Operations Management*, Vol. 27, No. 8, 1595-1610, doi: 10.1111/poms.12885.
[20] Kirch, M., Poenicke, O., Richter, K. (2017). RFID in logistics and production – applications, research and visions for smart logistics zones, *Procedia Engineering*, Vol. 178, 526-533, doi: 10.1016/j.proeng.2017.01.101.
[21] Golpîra, H., Khan, S.A.R., Safaeipour, S. (2021). A review of logistics Internet-of-Things: Current trends and scope for future research, *Journal of Industrial Information Integration*, Vol. 22, Article No. 100194, doi: 10.1016/j.jii.2020.100194.
[22] Issaoui, Y., Khiat, A., Bahnasse, A., Ouajji, H. (2019). Smart logistics: Study of the application of blockchain technology, *Procedia Computer Science*, Vol. 160, 266-271, doi: 10.1016/j.procs.2019.09.467.
[23] Ahmed, A.M., Sayar, A.R.Z. (2023). An empirical study on the factors influencing the implementation of blockchain-based supply chain traceability system, *Journal of Logistics, Informatics and Service Science*, Vol. 10, No. 1, 280-297, doi: 10.33168/LISS.2023.0116.
[24] Cavone, G., Dotoli, M., Seatzu, C. (2018). A survey on petri net models for freight logistics and transportation systems, *IEEE Transactions on Intelligent Transportation Systems,* Vol. 19, No. 6, 1795-1813, doi: 10.1109/TITS.2017.2737788.
[25] Jabeur, N., Al-Belushi, T., Mbarki, M., Gharrad, H. (2017). Toward leveraging smart logistics collaboration with a multi-agent system based solution, *Procedia Computer Science,* Vol. 109, 672-679, doi: 10.1016/j.procs.2017.05.374.
[26] Li, S., Sun, Q., Wu, W. (2019). Benefit distribution method of coastal port intelligent logistics supply chain under cloud computing, *Journal of Coastal Research,* Vol. 13, No. 93, 1041-1046, doi: 10.2112/SI93-150.1.

[27] 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.
[28] Barenji, A.V., Wang, W.M., Li, Z., Guerra-Zubiaga, D.A. (2019). Intelligent E-commerce logistics platform using hybrid agent based approach, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 126, 15-31, doi: 10.1016/j.tre.2019.04.002.
[29] Vukićević, A., Mladineo, M., Banduka, N., Mačužić, I. (2021). A smart Warehouse 4.0 approach for the pallet management using machine vision and Internet of Things (IoT): A real industrial case study, *Advances in Production Engineering & Management,* Vol. 16, No. 3, 297-306, doi: 10.14743/apem2021.3.401.
[30] Cao, G., Wang, Y., Gao, H., Liu, H., Liu, H., Song, Z., Fan, Y. (2023). Coordination decision-making for intelligent transformation of logistics services under capital constraint, *Sustainability,* Vol. 15, No. 6, Article No. 5421, doi: 10.3390/su15065421.
[31] Liu, W., Long, S., Liang, Y., Wang, J., Wei, S. (2023). The influence of leadership and smart level on the strategy choice of the smart logistics platform: A perspective of collaborative innovation participation, *Annals of Operations Research*, Vol. 324, 893-935, doi: 10.1007/s10479-021-04063-7.
[32] Liu, W., Hou, J., Yan, X., Tang, Q. (2021). Smart logistics transformation collaboration between manufacturers and logistics service providers: A supply chain contracting perspective, *Journal of Management Science and Engineering,* Vol. 6, No. 1, 25-52, doi: 10.1016/j.jmse.2021.02.007.
[33] Liu, W., Long, S., Wei, S., Xie, D., Wang, J., Liu, X. (2022). Smart logistics ecological cooperation with data sharing and platform empowerment: An examination with evolutionary game model, *International Journal of Production Research*, Vol. 60, No. 13, 4295-4315, doi: 10.1080/00207543.2021.1925173.
[34] Yan, Y., Zhao, R., Liu, Z. (2018). Strategic introduction of the marketplace channel under spillovers from online to offline sales, *European Journal of Operational Research*, Vol. 267, No. 1, 65-77, doi: 10.1016/j.ejor.2017.11.011.
[35] Zhang, S., Zhang, J. (2020). Agency selling or reselling: E-tailer information sharing with supplier offline entry, *European Journal of Operational Research,* Vol. 280, No.1, 134-151, doi: 10.1016/j.ejor.2019.07.003.
[36] He, P., He, Y., Xu, H., Zhou, L. (2019). Online selling mode choice and pricing in an O2O tourism supply chain considering corporate social responsibility, *Electronic Commerce Research and Applications*, Vol. 38, Article No. 100894, doi: 10.1016/j.elerap.2019.100894.
[37] Xu, X., He, P., Fan, Y. (2021). The pricing and carbon abatement decisions of a manufacturer selling with marketplace or reselling mode, *International Transactions in Operational Research*, Vol. 29, No. 2, 1220-1245, doi: 10.1111/itor.13025.
[38] Liu, Y., Ma, D., Hu, J., Zhang, Z. (2021). Sales mode selection of fresh food supply chain based on blockchain technology under different channel competition, *Computers & Industrial Engineering*, Vol. 162, Article No. 107730, doi: 10.1016/j.cie.2021.107730.
[39] Liu, H., Kou, X., Liu, H., Gao, H., Zhao, X. (2022). Which operating mode is the best? Consider different combinations of sales contracts and service methods, *Computers & Industrial Engineering,* Vol. 168, Article No. 108069, doi: 10.1016/j.cie.2022.108069.
[40] Geng, X., Tan, Y., Wei, L. (2018). How add-on pricing interacts with distribution contracts, *Production and Operations Management*, Vol. 27, No. 4, 605-623, doi: 10.1111/poms.12831.
[41] Kim, Y.J., Lee, W.-G. (2022). An online sales platform based design of limited edition production and sales systems, *Journal of Logistics, Informatics and Service Science,* Vol. 9, No. 2, 1-17, doi: 10.33168/LISS.2022.0201.
[42] Qin, X., Liu, Z., Tian, L. (2021). The optimal combination between selling mode and logistics service strategy in an e-commerce market, *European Journal of Operational Research*, Vol. 289, No. 2, 639-651, doi: 10.1016/j.ejor.2020.07.029.
[43] Li, D., Liu, Y., Fan, C., Hu, J., Chen, X. (2021). Logistics service strategies under different selling modes, *Computers & Industrial Engineering*, Vol. 162, Article No. 107684, doi: 10.1016/j.cie.2021.107684.
[44] Zhao, Y., Jia, W. (2021). Considering the sales method and logistics service strategy, which operating mode is the best under different power structures? *Mathematical Problems in Engineering,* Vol. 2021, Article ID 7653068, doi: 10.1155/2021/7653068.
[45] Liu, W.-H., Xu, X.-C., Kouhpaenejad, A. (2013). Deterministic approach to the fairest revenue-sharing coefficient in logistics service supply chain under the stochastic demand condition, *Computers & Industrial Engineering*, Vol. 66, No. 1, 41-52, doi: 10.1016/j.cie.2013.06.008.
[46] Yunmiao, G., Gong, B.G., Cheng, Y.M. (2009). Logistics service supply chain coordination under demand uncertainty, *Jisuanji Jicheng Zhizao Xitong/Computer Integrated Manufacturing Systems,* Vol. 15, No. 12, 2412-2416+2438.
[47] Liu, W., Liu, Y., Zhu, D., Wang, Y., Liang, Z. (2016). The influences of demand disruption on logistics service supply chain coordination: A comparison of three coordination modes, *International Journal of Production Economics*, Vol. 179, 59-76, doi: 10.1016/j.ijpe.2016.05.022.
[48] Wang, N., Fan, Z.-P., Wang, X. (2016). Channel coordination in logistics service supply chain considering fairness, *Mathematical Problems in Engineering*, Vol. 2016, Article ID 9621794, doi: 10.1155/2016/9621794.
[49] Liu, W., Wang, D., Shen, X., Yan, X., Wei, W. (2018). The impacts of distributional and peer-induced fairness concerns on the decision-making of order allocation in logistics service supply chain, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 116, 102-122, doi: 10.1016/j.tre.2018.05.006.
[50] Qin, X., Su, Q., Huang, S.H., Wiersma, U.J., Liu, M. (2019). Service quality coordination contracts for online shopping service supply chain with competing service providers: Integrating fairness and individual rationality, *Operational Research,* Vol. 19, 269-296, doi: 10.1007/s12351-016-0288-z.

[51] Niu, B., Dai, Z., Liu, Y., Jin, Y. (2022). The role of Physical Internet in building trackable and sustainable logistics service supply chains: A game analysis, *International Journal of Production Economics*, Vol. 247, Article No.108438, doi: 10.1016/j.ijpe.2022.108438.
[52] Kim, Y.-J., Ha, B.-C. (2022). Logistics service supply chain model, *Journal of Logistics, Informatics and Service Science,* Vol. 9, No. 3, 284-300, doi: 10.33168/LISS.2022.0320.

Appendix A

Proof of Corollary 1

(1) $p^{PC*} - p^{PS*} = \frac{az(-4-2u(-6+z)+z+u^2z)}{4(4-z)(4+(-1+u)z)}$, because $4 - z > 0$ and $4 + (-1 + u)z > 0$, then it only needs to satisfy that $-4 - 2u(-6 + z) + z + u^2z$ is positive，we can get $p^{PC*} > p^{PS*}$. We know that $s > 0$ when $\frac{-6-2\sqrt{9-2z}+z}{z} < u < 0.5$， then $-4 - 2u(-6 + z) + z + u^2z$ is always greater than 0，we can get $p^{PC*} > p^{PS*}$. At the same time $p^{PS*} - p^{PB*} = \frac{az(1-u)}{8-2z(1-u)} > 0$，we can get $p^{PB*} < p^{PS*} < p^{PC*}$. Proof relationship between q^{PB*}, q^{PS*} and q^{PC*}, and proof of relationship between y^{PS*} and y^{PC*} is similar to the proof of the relationship between p^{PB*}, p^{PS*} and p^{PC*}, here we omit.

(2)$\pi_T^{PC*} - \pi_T^{PS*} = \frac{a(-4z+12uz+z^2-2uz^2+u^2z^2)}{4(4-z)(4-z+uz)}$，because $4 - z > 0$ and $4 + (-1 + u)z > 0$，then it only needs to satisfy that $-4z + 12uz + z^2 - 2uz^2 + u^2z^2$ is positive，we can get $\pi_T^{PC*} > \pi_T^{PS*}$. Solve the quadratic equation of $-4z + 12uz + z^2 - 2uz^2 + u^2z^2 = 0$ with respect to u，we can get $\frac{-6-2\sqrt{9-2z}+z}{z} < u < 0.5$，at this time $-4z + 12uz + z^2 - 2uz^2 + u^2z^2$ is always greater than 0，we can get，so$\pi_T^{PC*} > \pi_T^{PS*}$. Meanwhile, $\pi_T^{PS*} - \pi_T^{PB*} = \frac{a(1-u)z}{2(4-(1-u)z} > 0$，we can get $\pi_T^{PB*} < \pi_T^{PS*} < \pi_T^{PC*}$. Proof relationship between π_J^{PB*}, π_J^{PC*} and π_J^{PS*} is similar to the proof of the relationship between π_T^{PB*}, π_T^{PS*} and π_T^{PC*}, here we omit.

Proof of Corollary 2.

The proof of Corollary 2 is similar to the proof of Corollary 1, here we omit.

Proof of Corollary 3.

The proof of Corollary 3 is similar to the proof of Corollary 1, here we omit.

Proof of Corollary 4.

The proof of Corollary 4 is similar to the proof of Corollary 1, here we omit.

Proof of Corollary 5.

The proof of Corollary 5 is similar to the proof of Corollary 1, here we omit.

APEM journal

Advances in Production Engineering & Management | **ISSN 1854-6250**
Volume 18 | Number 3 | September 2023 | pp 345–356 | Journal home: apem-journal.org
https://doi.org/10.14743/apem2023.3.477 | Original scientific paper

Factors affecting Quality 4.0 implementation in Czech, Slovak and Polish organizations: Preliminary research

Wawak, S.[a], Sütőová, A.[b], Vykydal, D.[c,*], Halfarová, P.[c]

[a]Department of Management Processes, Krakow University of Economics, Poland
[b]Faculty of Materials, Metallurgy and Recycling, Technical University of Košice, Slovakia
[c]Faculty of Materials Science and Technology, VSB -Technical University of Ostrava, Czech Republic

A B S T R A C T

With the proliferation of the Industry 4.0 paradigm, the inadequacy of conventional quality management tools has become increasingly apparent. The preliminary investigation presented in this paper focuses on the identification of the Quality 4.0 readiness level of organizations operating in the Czech Republic, Poland, and Slovakia, as well as affecting factors. The study is based on the review of relevant literature. The web-based questionnaire enabling organizations' representatives through Computer-Assisted Web Interviewing (CAWI) to take part in the study was used. Data from 298 completed responses were subjected to comprehensive analysis. Descriptive statistics and hypothesis testing were applied to analyse the data. Small and medium-sized organizations achieve low levels of Quality 4.0 readiness. Large organizations are better prepared. The study confirmed the dependence between the Quality 4.0 readiness level and whether the organization operates in automotive, while automotive organizations achieved a higher level of Quality 4.0 readiness than other organizations. The significant relationship between the Quality 4.0 readiness level and whether the organization has a certified management system was also confirmed. Received data also enabled the identification of the main barriers and benefits of Quality 4.0 implementation perceived by the organizations. The research findings identify the challenges that enterprises face regarding the Quality 4.0 implementation and the necessary support that organizations require. These findings can be a foundation for developing novel research initiatives and implementation programs. The research results contribute to the existing body of knowledge and bring new information and insights into the field of quality digitalization.

A R T I C L E I N F O

Keywords:
Quality management;
Industry 4.0;
Quality 4.0;
Quality 4.0 readiness;
Management systems;
Industry sector;
Organization size;
Chi-square test

**Corresponding author:*
david.vykydal@vsb.cz
(Vykydal, D.)

Article history:
Received 26 September 2023
Revised 25 October 2023
Accepted 27 October 2023

1. Introduction

During the latter decades of the 20th century and the initial years of the 21st century, there was a notable acceleration in technological development, accompanied by increasingly rapid changes in the conditions under which organizations functioned. Moreover, the emergence of significant possibilities for integrating new, ground-breaking technologies prompted practitioners and researchers not to speak of evolution but of a fourth industrial revolution. Key technologies that distinguish Industry 4.0 (I4.0) include Autonomous Robots (AR), Systems Integration (SI), Internet of Things (IoT), Cloud Computing (CC), Augmented Reality (AR), Big Data (BD), and simulations [1, 2].

Quality 4.0 (Q4.0) is an emerging research topic dealing with the question: How Quality Management (QM) needs to be adopted in the digital era? The term 'Quality 4.0' has emerged as the result of integrating I4.0 features with traditional QM practices. Q4.0 brings benefits for organizations like reduced costs of quality via reduced non-conformities and quality inspection, improved operational efficiencies, increased value proposition, transparent data-based partnership, and increased successful product and service innovations [3, 4]. There are only a few studies dealing with Q4.0 adoption in organizations, while most of them confirm a low level of Q4.0 readiness or maturity, e.g. [5-7]. The number of sources focusing on I4.0 readiness or maturity assessment is much higher and the problem is examined more in terms of the factors affecting the level of I4.0 maturity in organizations. Several I4.0 maturity models have been developed and applied [8-10]. Also, there are a few studies confirming differences in I4.0 readiness or maturity level depending on countries, e.g. [11, 12], size of the organization, e.g. [13-17], industry sector [18-20], etc. The problem of Q4.0 readiness level in organizations and factors contributing to the implementation of Q4.0 is little explored. Therefore, our paper focuses on the examination of Q4.0 readiness of organizations operating in the three Visegrad countries that belong among the most industrialized countries in the European Union – Poland, Czech Republic, and Slovakia. It examines factors that may relate to Q4.0 readiness level like organization size, industry sector, certified management systems in organizations and country of origin of organization These factors haven't been examined before in the context of Quality 4.0 readiness levels. It also focuses on the study of benefits and barriers of Q4. 0 implementation perceived by organizations. The research questions were defined as follows:

- What is the Q4.0 readiness of Czech, Polish and Slovak organizations?
- What are the main barriers and benefits of Q4.0 adoption perceived by organizations?
- Is there a significant relationship between organization size and Q4.0 readiness?
- Is there a significant relationship between Q4.0 readiness and whether the organizations have or don't have a certified quality management system (QMS)?
- Is there a significant relationship between Q4.0 readiness and whether organizations operate in the automotive industry or not?
- Is there a significant relationship between the country where organizations operate (Czech Republic, Poland, Slovakia) and Q4.0 readiness?

The findings of this study contribute to the existing body of knowledge by identifying Q4.0 readiness levels in organizations and affecting factors. The findings can help practitioners to understand the current state of transformation initiatives in this field and related aspects.

2. Literature background

2.1 Industry in Czech Republic, Poland and Slovakia and support of I4.0 on the level of countries

In the EU countries service sector employs most of the population. The Czech Republic, Slovakia and Poland are at the bottom of the ranking as far as employment in services is concerned. The countries belong among the six most industrialized economies in the EU. The Czech Republic is in second place after Ireland with the industry sharing 30.6 % of the Gross Domestic Product (GDP). The share of Poland is 29.8 % and the share in Slovakia is 28,6 % [21].

The most significant industry sector in the Czech Republic is the automotive industry with a 10 % share of GDP. Important manufacturers of passenger cars are Škoda Auto owned by the Volkswagen Group, Toyota, Peugeot Citroën Automobile and Hyundai. After the automotive, the chemical industry with a 7 % share of GDP followed by electrotechnics, machinery and metallurgy belong to the most important industry sectors. As in the Czech Republic, the automotive industry is the most important sector in Slovakia with a 13.9 % share of GDP. It accounts for 47 % of total industry production. Currently, four car makers are operating in Slovakia – VW, Stellantis, Kia, and Jaguar Land Rover. Slovakia is the world leader in car production per capita. Other high-value-added industries are the chemical industry (10 % share of the total industrial production), electronics and electrical components (9.3 % share of the total industry share), machinery, metallurgy and metal proceeding industry [22]. In comparison with the Czech Republic and Slovakia, Poland's reliance on the automotive industry is lower. It represents just 3.4 % of GDP [23]. Significant industry sectors in Poland are the mechanical and electromechanical industry followed by the food industry, metallurgical and chemical industry [24].

Digital transformation is inevitable to maintain the countries' economic competitiveness. The latest results of the European Innovation Scoreboard show that the Czech Republic belong to the moderate innovator and Slovakia and Poland to the group of emerging innovators [25]. The Digital Economy and Society Index (DESI) ranked the Czech Republic in 19th place, Slovakia in 23rd and Poland in 24th place [26]. According to the World Digital Competitiveness ranking Czech Republic took 33rd place, Poland 46th and Slovakia 47th from a total of 64 countries [27]. The Czech Republic in comparison with Poland and Slovakia achieved the highest ranking on the base of the above-mentioned studies. According to the survey by the European Investment Bank, 79 % of firms in Slovakia use advanced digital technologies, 72 % in the Czech Republic and 66 % in Poland while the EU average is 69 %. As the industrial sector is important for the economy of the countries, there is an interest in supporting digital transformation. Strategic initiatives like the National Industry 4.0 Initiative in the Czech Republic (2015) and the Concept of Smart Industry in Slovakia (2016) were approved. In Poland, no Industry 4.0 individual strategic document was developed but the Future Industry Platform was established in 2019 as a part of the Responsible Development Plan to create mechanisms for cooperation and interdisciplinary knowledge transfer for accelerating digital transformation. National Industry 4.0 platforms were founded also in Slovakia (Smart Industry Platform) and the Czech Republic (National Centre for Industry 4.0). There have been established several Digital Innovation Hubs (DIHs) and European Digital Innovation Hubs (EDIHs) in the countries. Digital transformation of industries is also supported by cross-sectional strategies supplementing the above-mentioned initiatives like Digital Czech Republic (2018), the Digital Transformation Strategy of the Slovak Republic (2018) and Poland's Strategy of Responsible Development (2017) incorporating Industry 4.0 problematic.

2.2 Quality 4.0 and Quality 4.0 maturity and readiness models

Q4.0 is a relatively new term that has emerged in relation to I4.0. The field of Quality Management is essential for ensuring the required quality of products and services and customer satisfaction. Approaches to quality have gone through several development stages from Quality Inspection through Quality Control, Quality Assurance to Total Quality Management and now the era of I4.0 is forcing the development of the existing approaches towards Q4.0. Q4.0 as an emerging concept representing the next developmental stage of QM has attracted much attention from scholars, practitioners as well as consulting organizations (e.g. BCG, The Oakland Group, Juran Institute) during the last years. After the review of papers in the Web of Science (WoS) database containing the term Q4.0 in the title or abstract, 98 publications were found, while 15 of them were eliminated as they did not relate to the Q4.0 as well as the other 5 papers, that mentioned Q4.0 only by few words and the presented studies dealt with another area. The focus of the remaining 78 publications published from 2016 to 2023 is presented in Table 1.

Table 1 Focus of Q4.0 publications indexed in the WoS database

Publications No.	Focus
22	Q4.0 definition, and/or Q4.0 principles and characteristics and/or Q4.0 advantages, disadvantages
20	Selected Q4.0 tools (implementations or review of selected tools and discussion)
14	Q4.0 maturity or readiness assessment or assessment of usage level of Q4.0 technologies
11	Identification of main determinants, dimensions of Q4.0 and Q4.0 framework development
5	Q4.0 competencies and/or relation between Q4.0 and human factor or Leadership's impact on Q4.0 implementation
3	Lean approaches in connection with Q4.0
2	Q4.0 in relation to sustainability or circular economy
1	Q4.0 impact on organizational performance

Q4.0 as a term was described in several publications, however, a uniform definition has not been established so far. According to the American Association for Quality, Q4.0 references organizational excellence within the context of I4.0 [28]. Q4.0 aligns quality management with I4.0, which results in increased efficiency, performance and improved business models. Q4.0 uses new technologies like BD, IoT and AI with existing quality methods to broaden the scope of QM and deal with a completely new set of complex problems. Some other examples of Q4.0 tools and methods include digital twin technology, which enables the creation of virtual models of products and processes, enabling simulation and optimization as well as blockchain technology, which allows secure and transparent tracking of supply chains and product histories. Using advanced technologies helps to design, operate and maintain predictive, adaptive, automated quality systems along with improved human interaction through quality planning, assurance and improvement to achieve new optimums in performance, operational excellence, and innovation. Q4.0 emphasizes the integration of QM to ensure a holistic approach to quality throughout the entire value chain. Researchers endeavour to define Q4.0 by highlighting its distinctive features. They note that this is a concept that promotes the adoption of contemporary QM methods, which are grounded, among other things in:

- customer value co-creation enabled by vertical and horizontal integration,
- cross-functional collaboration,
- eliminated visual and manual inspection,
- human empowerment and human-robot interaction,
- integration of the organisation's physical infrastructure and processes with the network and databases,
- collecting and analysing live data on the functioning of the infrastructure and processes,
- fast, adaptive learning and introducing changes before problems occur (prediction),
- using ML and AI for monitoring, analysis, and fast decision-making,
- improved trust, transparency, and auditability.

There are a few Q4.0 readiness or maturity models that have been published so far. Table 2 presents these models while elements of individual models were assigned to the selected areas – governance and culture, processes, people, technology, and results. Many of the dimensions defined by the models and related elements overlap but they are named differently. For that reason, the elements related to the dimensions of the models were assigned to the above-mentioned areas. The Q4.0 model published by LSN involving 11 elements was the first published framework in this field. It helps to interpret the organization's current state and identify what changes need to be done to move towards Q4.0. The transformational levels are defined for every element. In other cases of Q4.0 maturity models the area of Process often involves elements that cover Q 4.0 technologies used for process control. The three readiness models in Table 2 define the prerequisites for Quality 4.0 and focus mainly on the first three areas. The model published by [29, 30] defines the certified QMS as a prerequisite for successful Q4.0 implementation.

Among the challenges related to Q4.0 in terms of its implementation, management commitment to invest in technology and missing Industry 4.0 strategy of the organization were identified as the most important [5]. The study conducted by [36] among the top challenges identified the high cost of implementation, lack of resources, lack of knowledge, organization culture and not clear competitive advantage offered by Q4.0. The motivation factors for Q4.0 implementation involved accessibility of information, BD-driven QM programs, increased customer satisfaction, productivity improvement and cost saving [37, 38].

The number of studies dealing with the factors related to I4.0 maturity or readiness level is much higher. Several studies confirmed the relationship between Industry 4.0 maturity level and size of the organization [13-17] and industry sectors [18-20], while among the most matured sectors were the automotive, electronics and pharmaceutical industries.

In our study, we focused on the examination of selected factors in relation to the Q4.0 readiness levels as well as motivators driving Q4.0 implementation in organizations and main barriers avoiding the digital transformation of the traditional approaches to quality.

Table 2 Q4.0 publications focus in the WoS database

Q4.0 Maturity/Readiness Model	No. of items within the areas					Sum of Elements (∑)	Maturity/ Readiness level
	Governance, culture	Processes	People	Technology	Results		
Q4.0 Transformation Model [32]	2	3	2	4	-	11	Levels for each element
Q4.0 Maturity Model [7]	4	4	5	9	-	22	7 Q4.0 maturity levels
Q4.0 Maturity Index [33]	4	12	4	8	-	28	5 Q4.0 maturity levels
Q Intelligence Maturity Model [34]	1	4	-	2	1	8	5 Stages of Q maturity ladder
Q Intelligence Maturity Index [35]	2	8	1	2	1	14	4 Q intelligence maturity stages
Q4.0 Readiness Assessment Tool [31]	16	6	3	-	-	25	5 Q4.0 readiness levels
Q4.0 Readiness Assessment Tool [36]	4	2	2	-	-	8	5 Q4.0 readiness levels
Q4.0 Readiness Assessment Tool [6]	5	2	3	2	-	12	5 Q4.0 readiness levels

3. Methodology

Conducting preliminary research is essential to establish familiarity with the phenomenon under study, determine the importance and intensity of its features, and identify factors that may significantly impact the research outcome. This preliminary research serves as a foundation for obtaining valuable initial knowledge about the subject of investigation and concurrently highlights areas that require further exploration and development. Preliminary research can be carried out using formalized and structured methods and unstructured methods with a low level of formalization. To achieve the goals of the study, quantitative methods were used. For data collection online questionnaire was developed with closed multiple-choice questions. The questionnaire contained items focusing on:

- segmentation characteristics of organizations (size, industry type),
- types of implemented management systems in organizations,
- benefits and barriers of Q4.0 implementation perceived by respondents,
- Q4.0 readiness level in organizations.

The questionnaires were distributed through a dedicated internet portal (CAWI) to organizations of different sizes and sectors operating in the Czech Republic, Poland and Slovakia. Additionally, information about the study was disseminated through professional social networking sites. Data collection was carried out between April and July 2022. The questionnaires were filled out by quality managers or integrated management system representatives of the organizations.

Descriptive statistics and hypothesis testing were used to evaluate the data collected through the questionnaires. For hypothesis testing the Chi-square test was used, while the following hypotheses were proposed:

- H1: There is a dependence between the Quality 4.0 readiness level and the size of the organization.
- H2: There is a dependence between the Quality 4.0 readiness level and whether the organisation operates in the automotive industry.
- H3: There is a dependence between the Q4.0 readiness level and the country in which the organization operates.
- H4: There is a dependence between the level of Q4.0 readiness and whether the organisation has a certified quality management system or doesn't have.

The null hypothesis H0 acceptance and rejection of alternative hypothesis H1 in the case of the above-mentioned hypotheses confirms the significant dependence between the examined parameters. Otherwise, it confirms that there isn't a significant relationship between the parameters. For evaluation of Quality, 4.0 readiness 6 levels were used:

- Level 0 – the organization is not prepared for Q4.0 at all.
- Level 1 – information and automation technologies are used isolated without mutual connection.
- Level 2 – information systems and infrastructure elements are connected to the network but without the possibility of control of processes in real-time.
- Level 3 – digitalization enables real-time control of processes and communication.
- Level 4 – big data from internal processes and external processes are analysed to predict future state.
- Level 5 – decisions are realized automatically through intelligent systems that are widely used in organizations.

On the base of the literature review, we assumed that a significant proportion of the surveyed organizations are likely to be in the early stages of implementing the Quality 4.0 concept. Completed implementation projects in the Quality 4.0 domain are relatively scarce among organizations. Nevertheless, there is a growing interest among organizations in this area. Evaluating the actual level of preparedness, perceived benefits, and obstacles and the relation of the selected factors with the level of Q4.0 readiness will contribute to the existing body of knowledge and provide valuable information for practitioners.

4. Results and discussion

4.1 Research sample

Totally 298 questionnaires were received, 121 in the Czech Republic, 101 in Poland and 76 in Slovakia. The received questionnaires represented over 20 industry sectors. The most represented industries were the automotive industry (27 %), mechanical industry (16 %) chemical industry and plastics processing (12 %). The structure of respondents by industry sector is presented in Fig. 1.

Medium and large organizations dominated the study. Specifically, 20 % of the respondents represented organizations with over 1,000 employees. Meanwhile, 29 % of the respondents were from large organizations, and 34 % were from medium-sized organizations. The breakdown of participating organizations according to their size in terms of the number of employees is shown in Fig. 2.

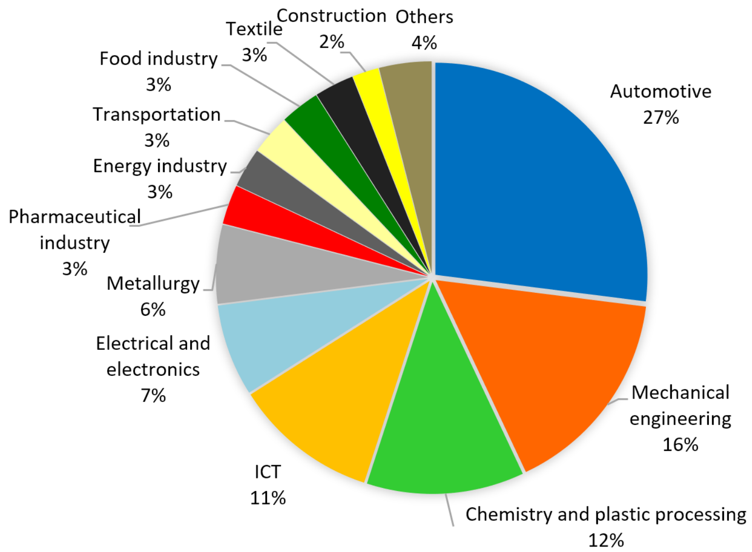


Fig. 1 Organizations by industry sector

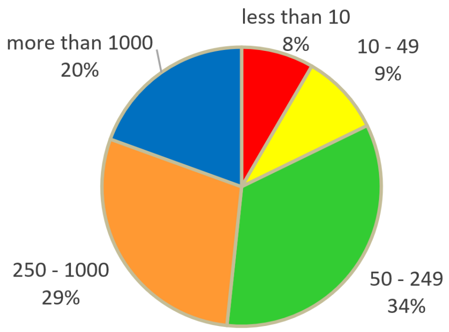


Fig. 2 Organizations according to the number of employees

According to the results, 54 % of the respondents reported that their organization have a certified ISO 9001 QMS, while 30 % indicated an integrated management system (IMS) involving environmental or occupational health and safety management system. As many as 7 % of respondents declared the implementation of the ISO 13485 system and 29 % of IATF 16949. Additionally, 19 % of respondents declared implementation and certification of at least one MS and 17 % of organizations didn't have any certified management system. Fig. 3 illustrates the percentage of organizations with individual management systems. The respondents could choose more options.

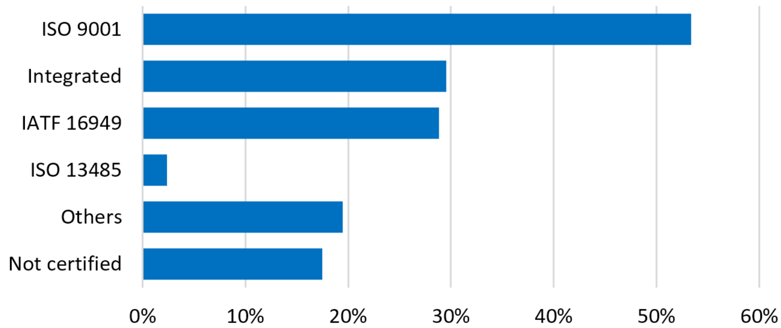


Fig. 3 Implemented management systems in organizations

4.2 Potential benefits and barriers of Q 4.0 implementation

Respondents were questioned about their perception of the main benefits offered by Q4.0 implementation while multiple responses were offered. The results are presented in Fig. 4.

The three most important benefits listed by the respondents were the creation of conditions for long-term ability to succeed in a competitive environment (52 %), support of interconnection of processes and levels of management (48 %), performance increase of all processes (44 %). Organizations with an IMS or IATF 16949 consider support of interconnection of processes more often as a significant benefit than those with only ISO 9001 or without these systems. The benefit of competitive advantage was confirmed by organizations with ISO 9001 or IATF (64 % and 73 %). Mass customisation of products (12 %) and the possibility of achieving compliance at the 6-sigma level (12 %) are considered the least significant benefits of Q4.0 implementation. Almost 67 % of very large organizations consider achieving the 6 Sigma level as a benefit, and 59 % the increased flexibility of interventions in case of product deviations and process specifications.

The most significant barriers of Q4.0 implementation perceived by the respondents are shown in Fig. 5.

Time and investment requirements are the most significant barriers considered by 70 % of organizations, followed by the current lack of financial resources (35 %) and the absence of a long-term QM strategy. Organisations without certificated management systems more often emphasised the need to supplement knowledge (39 %) as a barrier of Q4.0 implementation than those with certified management systems.

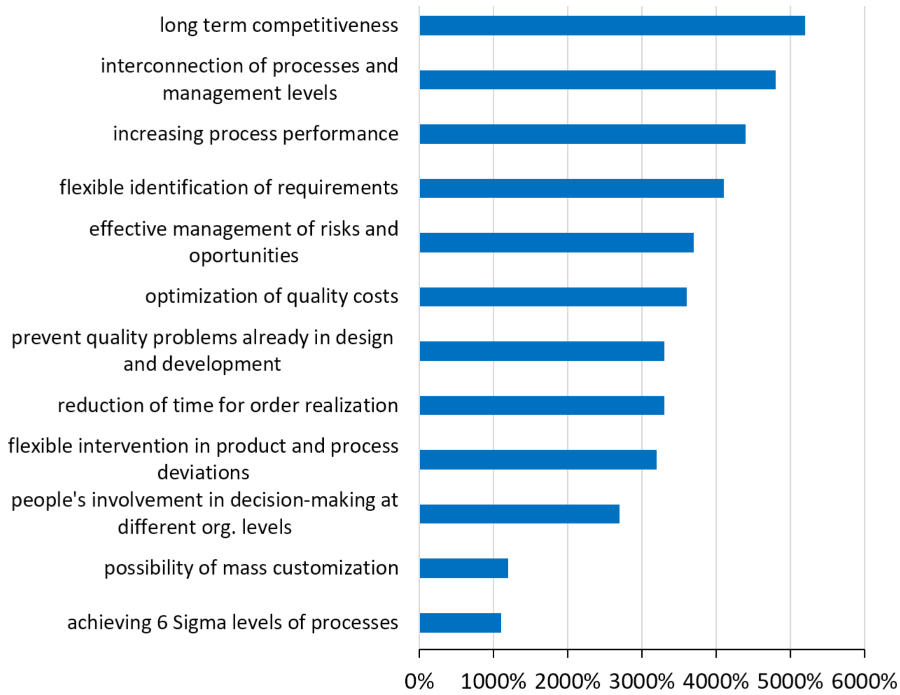


Fig. 4 Potential benefits of Q4.0 implementation

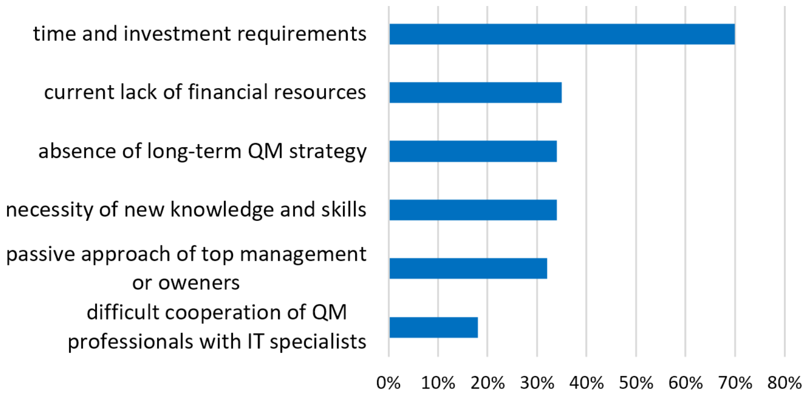


Fig. 5 Potential barriers of Q4.0 implementation

4.3 Quality 4.0 readiness level

Almost 22 % of respondents stated that their organization is not ready to implement Q4.0 (level M0). Level M1 characterized by isolated automation and information systems was chosen by 15 % of organizations. The level M2 representing partially interconnected information systems was typical for 16 % of organizations. Level M3 described by connected information systems and infrastructure without the possibility of process control in real-time is achieved by 12 % of organizations. Level M4 achieved the second highest value (20 %) and represents organizations where digitalization enables the control of processes in real-time. There is a low percentage of organizations achieving level M5 (10 %) where advanced analytics is used for proceeding the big data to make predictions and the highest level M6 characterized by the possibility of automated decision-making enabled by intelligent technologies based on big data and advanced analytics is typical for 6.5 % of organizations. Fig. 6 shows the percentage value of the organization classified into individual Q4.0 readiness levels.

In micro and small organizations prevail the M0 and M1 levels. In the case of medium-sized organizations, the Q4.0 readiness level rises, but only 27 % of organizations achieve M4, M5 or M6 levels. More than 40 % of large organizations achieve the three highest levels and in the case of very large organizations, it is more than 60 % of organizations.

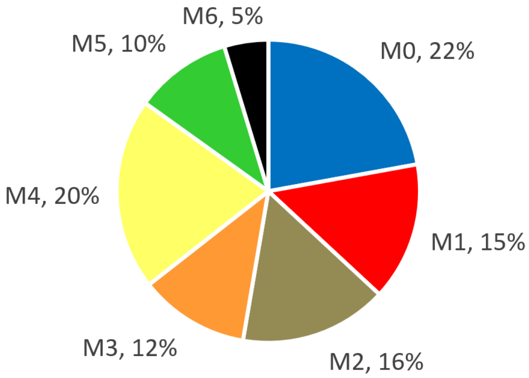


Fig. 6 Quality 4.0 readi ness levels

Fig. 7 shows the percentage of organizations of different sizes within individual Q4.0 readiness levels.

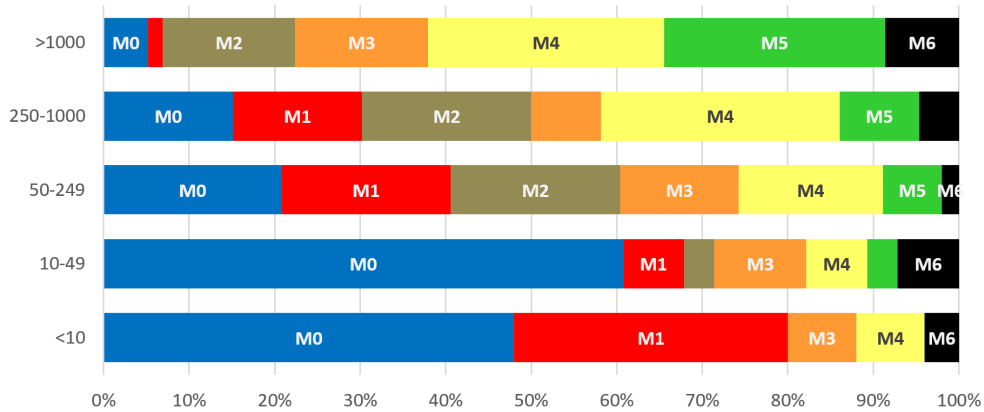


Fig. 7 Quality 4.0 readiness levels in the organizations with different numbers of employees

4.4 Hypotheses testing

From the results in Fig. 7 it can be concluded that the level of readiness for Q.4 implementation is related to the size of the organisations. The results of the H1 testing (p-value = 0.000) confirm that there is a significant statistical dependence between the level of organizational readiness to implement Q4.0 and the size of the organization. The test used does not allow to define what is the cause and the effect, however, in the context of Fig. 7 it can be concluded that large organisations are better prepared for Q.4 implementation than small and medium organizations. The fact that I4.0 maturity and readiness levels are higher in large organizations than small and medium-sized was confirmed in some studies, e.g. [13-17].

The results of the hypothesis H2 testing (p-value = 0.002) show that there is a significant statistical dependence between the level of readiness for Q4.0 implementation and whether the organisation operate in the automotive industry. The organisations operating in the automotive industry have a higher level of Q4.0 readiness (average level is 4.221, i.e. readiness level between M3 and M4) than organisations from other sectors (average level is 2.795, i.e. readiness level between M1 and M2). This is in accordance with the percentages, where the results show that 54 % of non-automotive organisations ranked themselves to be at the Q4.0 readiness levels M0 and M1. In contrast, only 14 % of organisations from the automotive sector ranked themselves to be at the first two lowest readiness levels and 53.5 % of organisations ranked themselves to be at the highest levels M4 to M6. Some studies focusing on the Industry 4.0 readiness and maturity confirmed that automotive sector belongs to front runners, e.g. [18-20]. It is consistent with our finding.

On the basis of the results of the hypothesis H3 testing (p-value 0.001), it is possible to accept the null hypothesis and reject the alternative hypothesis, what means that there is a significant dependence between the country where the organization operates and the Quality 4.0 readiness level. This finding may be related to the different industry structures in the countries. In the Czech

Republic and Slovakia, in contrast to Poland, the automotive industry is predominant, where there are higher levels of Q4.0 readiness. It also can relate to different levels of digitalization in the countries confirmed by the studies like, e.g. [25-27].

Based on the H4 hypothesis testing (p-value = 0.002), it was concluded that there is a significant statistical dependence between the level of Q4.0 readiness level and whether the organisation has a certified QMS or doesn't have. It can be concluded that organisations with a certified QMS are better prepared for the implementation of Q4.0 (average level is 3.403, i.e. readiness levels between M2 and M3) than organisations without a certified QMS (average level is 2.736, i.e. readiness levels between M1 and M2). Again, the percentage share showed an interesting result. The lowest level of readiness was in the case of 49.1 % of organisations that do not have a certified QMS. Only 18.2 % of organizations with a certified QMS evaluated themselves to be at the M0 level.

5. Conclusion

Organizations that apply I4.0 technologies are experiencing technological advancements that reveal the limitations of current quality management tools. Implementation of advanced information technology systems equips quality experts with sophisticated data, necessitating their adept interpretation. Consequently, novel and updated quality tools must be developed, and new competencies must be defined and guaranteed.

During our preliminary research we focused on the Quality 4.0 readiness level of organizations operating in the selected Visegrad countries – Poland, Czech Republic and Slovakia. The results revealed that small and medium-sized organizations achieve low levels of Quality 4.0 readiness. Large organizations are better prepared. There are only a few studies focusing on Quality 4.0 readiness level that have been published confirming that organizations are in the early stages, what is consistent with our findings in terms of small and medium sized organizations (SMEs). Our results confirmed statistically significant dependence between the size of the organization and Quality 4.0 readiness. It was also confirmed the dependence between Quality 4.0 readiness and whether the organization operates in automotive or not. Automotive organizations achieved a higher level of Industry 4.0 readiness. Among the three main barriers of Quality 4.0 implementation perceived by organizations the investment requirements, the current lack of financial resources and the absence of a long-term QM strategy were identified. On the other hands, the organizations consider as three main benefits of Quality 4.0 implementation the long-term competitiveness, interconnection of processes and organization levels and increasing process performance. The study confirmed the dependence between Quality 4.0 readiness and the countries where the organizations operate. Organizations with certified (QMS) achieved higher levels of Q4.0 readiness and it was confirmed that there is a statistically significant dependence between Q4.0 readiness and whether the organization has implemented certified QMS.

Q4.0 is a recently developed concept and research in this area is in its early stages. The research findings identify the challenges that enterprises face regarding the Quality 4.0 implementation and the necessary support that they require. These findings can be a foundation for developing novel research initiatives and implementation programs. They can serve as an input for preparation of supporting initiatives for SMEs on the level of countries. The research results contribute to the existing body of knowledge and bring new information and insights into the field of quality digitalization and factors contributing to the transformation of traditional quality approaches for the needs of Industry 4.0 and can help organization to build suitable strategies The research conducted in this study is preliminary. A limitation of the study is the questionnaire's length. The examined factor - country of organization's origin in relation to Quality 4.0 readiness level must be further analysed. Also not all possible factors and dependencies were detected. However, these limitations can be addressed in future, more in-depth research.

Acknowledgement

The research and the publication were co-financed by a subsidy granted to the Krakow University of Economics – Project No. 049/ZZP/2023/POT. It was also supported by specific research projects No. SP2023/043 "Development of Integrated Management of Industrial Processes through Linking The Principles of Smart Production and Quality Management" as part of the Faculty of Materials Science and Technology, VŠB-TU Ostrava, with the support of the Ministry of Education, Youth and Sports, Czech Republic.

References

[1] Reagan, J.R., Singh, M. (2020). *Management 4.0, Cases and methods for the 4th industrial revolution*, Springer, New York, USA, doi: 10.1007/978-981-15-6751-3.

[2] Sader, S. Husti, I., Daróczi, M. (2019). Industry 4.0 as a key enabler toward successful implementation of total quality management practices, *Periodica Polytechnica Social and Management Sciences,* Vol. 27, No. 2, 131-140, doi: 10.3311/PPso.12675.

[3] Antony, J., McDermott, O., Sony, M. (2022). Quality 4.0 conceptualisation and theoretical understanding: A global exploratory qualitative study, *The TQM Journal,* Vol. 34, No. 5, 1169-1188, doi: 10.1108/TQM-07-2021-0215.

[4] Javaid, M., Haleem, A., Pratap Singh, R., Suman, R. (2021). Significance of Quality 4.0 towards comprehensive enhancement in the manufacturing sector, *Sensors International*, Vol. 2, Article No. 100109, doi: 10.1016/j.sintl.2021.100109.

[5] Antony, J., Sony, M., McDermott, O., Jayaraman, R., Flynn, D. (2023). An exploration of organizational readiness factors for Quality 4.0: An intercontinental study and future research directions, *International Journal of Quality & Reliability Management,* Vol. 40, No. 2, 582-606, doi: 10.1108/IJQRM-10-2021-0357.

[6] Maganga, D.P., Taifa, I.W.R. (2023). The readiness of manufacturing industries to transit to Quality 4.0, *International Journal of Quality & Reliability Management*, Vol. 40, No. 7, 1729-1752, doi: 10.1108/IJQRM-05-2022-0148.

[7] Nenadál, J., Vykydal, D., Halfarová, P., Tylečková, E. (2022). Quality 4.0 maturity assessment in light of the current situation in the Czech Republic, *Sustainability*, Vol 14, No. 12, Article No. 7519, doi: 10.3390/su14127519.

[8] Santos, R.C., Martinho, J.L. (2020). An Industry 4.0 maturity model proposal, *Journal of Manufacturing Technology Management,* Vol. 31, No. 5, 1023-1043, doi: 10.1108/JMTM-09-2018-0284.

[9] Schumacher, A., Erol, S., Sihn, W. (2016). A maturity model for assessing Industry 4.0 readiness and maturity of manufacturing enterprises, *Procedia CIRP*, Vol. 52, 161-166, doi: 10.1016/j.procir.2016.07.040.

[10] Wagire, A.A., Joshi, R., Rathore, A.P.S., Jain, R. (2021). Development of maturity model for assessing the implementation of Industry 4.0: Learning from theory and practice, *Production Planning & Control*, Vol. 32, No. 8, 603-622, doi: 10.1080/09537287.2020.1744763.

[11] Dujin, A., Geissler, C., Horstkötter, D. (2014). *Industry 4.0, The new industrial revolution, How Europe will succeed,* Roland Berger Strategy Consultants, Munich, Germany.

[12] World Economic Forum. (2018). *Readiness for the future of production, Report 2018*, World Economic Forum, Geneva, Switzerland.

[13] Ali, K., Johl, S.K. (2023). Impact of total quality management on Industry 4.0 readiness and practices: Does firm size matter?, *International Journal of Computer Integrated Manufacturing*, Vol. 36, No. 4, 567-589, doi: 10.1080/0951192X.2022.2128213.

[14] Frank, A.G., Dalenogare, L.S., Ayala, N.F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies, *International Journal of Production Economics*, Vol. 210, 15-26, doi: 10.1016/j.ijpe.2019.01.004.

[15] Grzyb, K. (2019). Industry 4.0 market in Poland from the international perspective, In: *Proceedings of International Scientific Conference Hradec Economic Days 2019*, Hradec Králové, Czech Republic, 252-262, doi: 10.36689/uhk/hed/2019-01-025.

[16] Moura, L.R., Kohl, H. (2020). Maturity assessment in Industry 4.0 – A comparative analysis of Brazilian and German companies, *Emerging Science Journal,* Vol. 4, No. 5, 365-375, doi: 10.28991/esj-2020-01237.

[17] Sundberg, L., Gidlund, K.L., Olsson, L. (2019). Towards Industry 4.0? Digital maturity of the manufacturing industry in a Swedish region, In*: Proceedings of 2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM),* Macao, China, 731-735, doi: 10.1109/IEEM44572.2019.8978681.

[18] Bittighofer, D., Dust, M. Irslinger, A., Liebich, M., Martin, L. (2018). State of Industry 4.0 across German companies, In: *Proceedings of 2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, Stuttgart, Germany, 1-8, doi: 10.1109/ICE.2018.8436246.

[19] Ozbiltekin-Pala, M., Kazancoglu, Y., Kumar, A., Garza-Reyes, J.A., Luthra, S. (2022). Analyzing critical factors of strategic alignment between operational excellence and Industry 4.0 technologies in smart manufacturing, *The TQM Journal,* doi: 10.1108/TQM-07-2022-0232.

[20] Závadská, Z., Závadský, J. (2020). Quality managers and their future technological expectations related to Industry 4.0, *Total Quality Management & Business Excellence*, Vol. 31, No. 7-8, 717-741, doi: 10.1080/14783363.2018.1444474.

[21] The World Bank. Industry (including construction), value added (% of GDP) - European Union, from *https://data.worldbank.org/indicator/NV.IND.TOTL.ZS?locations=EU&most_recent_value_desc=true*, accessed April 2, 2023.

[22] Sario, Slovak Investment and Trade Development Agency. Industries and Regions Overviews, from *https://www.sario.sk/en/invest-slovakia/industries-regions-overviews*, accessed May 7, 2023.

[23] EMIS. Poland automotive sector report 2022-2023, from *https://info.emis.com/poland-automotive-sector-2022-2023-report*, accessed May 20, 2023.
[24] Kulas, B. Przemysł w Polsce, from *https://geografia24.pl/przemysl-w-polsce/*, accessed June 6, 2023.
[25] European Commission. European innovation scoreboard, from *https://research-and-innovation.ec.europa.eu/statistics/performance-indicators/european-innovation-scoreboard_en*, accessed June 6, 2023.
[26] European Commission. DESI 2022, from *https://digital-strategy.ec.europa.eu/sk/node/11127*, accessed June 20, 2023.
[27] IMD. World Digital Competitiveness Ranking, from *https://www.imd.org/centers/wcc/world-competitiveness-center/rankings/world-digital-competitiveness-ranking*, accessed June 20, 2023.
[28] ASQ. Quality 4.0, from *https://asq.org/quality-resources/quality-4-0*, accessed July 4, 2023.
[29] Sony, M., Antony, J., Douglas, J.A. (2020). Essential ingredients for the implementation of Quality 4.0: A narrative review of literature and future directions for research, *The TQM Journal*, Vol. 32, No. 4, 779-793, doi: 10.1108/TQM-12-2019-0275.
[30] Pacana, A., Siwiec, D. (2021). Analysis of the possibility of used of the quality management techniques with non-destructive testing, *Tehnički Vjesnik – Technical Gazette,* Vol. 28, No. 1, 45-51, doi: 10.17559/TV-20190714075651.
[31] Zulfiqar, M., Antony, J., Swarnakar, V., Sony, M., Jayaraman, J., McDermott O. (2023). A readiness assessment of Quality 4.0 in packaging companies: An empirical investigation, *Total Quality Management & Business Excellence*, Vol 34, No. 11-12, 1334-1352, doi: 10.1080/14783363.2023.2170223.
[32] LSN Research. Quality 4.0 Impact and strategy handbook: Getting digitally connected to transform quality management, from *https://blog.lnsresearch.com/quality40ebookthank-you?submissionGuid=9ede63e5-aab1-4ef8-83f4-666d95481d36*, accessed May 12, 2023.
[33] Mtotywa, M.M. (2022). Developing a Quality 4.0 maturity index for improved business operational efficiency and performance, *Quality Innovation Prosperity,* Vol. 26, No. 2, 101-127, doi: 10.12776/qip.v26i2.1718.
[34] AlisQI. How to expose weaknesses in your Quality Management, from *https://www.alisqi.com/en/resources/whitepapers/how-to-expose-weaknesses-in-your-quality-management*, accessed July 7, 2023.
[35] Zonnenshain, A., Kenett, R.S.. (2020). Quality 4.0 – the challenging future of quality engineering, *Quality Engineering*, Vol. 32, No. 4, 614-626, doi: 10.1080/08982112.2019.1706744.
[36] Sony, M., Antony, J., Douglas, J.A., McDermott, O. (2021). Motivations, barriers and readiness factors for Quality 4.0 implementation: An exploratory study, *The TQM Journal,* Vol. 33, No. 6, 1502-1515, doi: 10.1108/TQM-11-2020-0272.
[37] Lalić, B., Delić, M., Simeunović, N., Tasić, N., Cvetković, S. (2019). The impact of quality management purchasing practices on purchasing performance in transitional economies, *Tehnički Vjesnik – Technical Gazette,* Vol. 26, No. 3, 815-822, doi: 10.17559/TV-20171130164306.
[38] Tian, S., Zhang, Z., Xie, X., Yu, C. (2022). A new approach for quality prediction and control of multistage production and manufacturing process based on Big Data analysis and Neural Networks, *Advances in Production Engineering & Management,* Vol. 17, No. 3, 326-338, doi: 10.14743/apem2022.3.439.

APEM journal
Advances in Production Engineering & Management | **ISSN 1854-6250**
Volume 18 | Number 3 | September 2023 | pp 357–370 | Journal home: apem-journal.org
https://doi.org/10.14743/apem2023.3.478 | Original scientific paper

Project portfolio management in telecommunication company: A stage-gate approach for effective portfolio governance

Milenkovic, M.[a], Ciric Lalic, D.[b,*], Vujicic, M.[c], Pesko, I.[b], Savkovic, M.[b], Gracanin, D.[b]

[a]Telecom Serbia, Belgrade, Republic of Serbia
[b]Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Republic of Serbia
[c]Faculty of Sciences, University of Novi Sad, Novi Sad, Republic of Serbia

A B S T R A C T

In today's fast-paced business environment, implementing strategies through programs, projects, and business-as-usual activities can be challenging for companies. The telecommunication industry, in particular, faces these challenges as it experiences the effects of digital transformation and fast-changing markets. It requires a flexible and adaptive approach to project portfolio management (PPM) to optimize investments and deliver value. This article presents a successful case study of a PPM process using the Stage-Gate model in a prominent telecommunications company that operates in a dynamic and fast-growing environment. The Stage-Gate PPM model comprises four stages: Proposal Selection, Selection of Nominated Demands, Prioritization, and Categorization of Projects. The model is unique as it can be adapted to different projects and incorporates elements of Agile approaches, such as Portfolio Sprint meetings and artefacts. The study demonstrates the importance of a well-defined PPM process in coordinating short-term and long-term activities and effectively allocating time, money, and resources. The Stage-Gate PPM model can potentially enhance project success rates and bring greater value to companies by ensuring the realization of suitable projects. This article contributes significantly to the existing literature on portfolio management, providing valuable insights and lessons applicable to other companies in the industry to enhance their portfolio management processes. Furthermore, this study can interest scholars and researchers seeking to explore effective portfolio management in other complex and dynamic environments.

A R T I C L E I N F O

Keywords:
Telecommunication industry;
Project portfolio management;
Stage-gate model;
Strategic goals;
Value delivery system

**Corresponding author:*
danijela.ciric@uns.ac.rs
(Ciric Lalic, D.)

Article history:
Received 30 May 2023
Revised 8 June 2023
Accepted 17 June 2023

1. Introduction

The telecommunications industry is constantly evolving, with new trends emerging as a technology and consumer preferences change. As stated in the *Telecom Services Market Size, Share & Trends Analysis Report By Service Type*, the telecommunications services market is expected to grow significantly in the coming years, driven by the increasing demand for mobile data services and the growth of machine-to-machine services [1]. In today's dynamic and highly competitive business environment, where customer demands are constantly evolving and competition is fierce, every industry faces significant challenges [2, 3]. Telecommunication companies, in particular, must possess the capability to adapt to these changes and uphold flexible internal processes in order to achieve success.

In the telecommunications industry, companies are process and project-oriented, relying on project management to solve business problems and adapt to the dynamic environment [4]. Environmental uncertainties and changing customer demands are just another factor pushing companies to be more flexible [5]. However, effective portfolio management techniques are essential to remain competitive and achieve their business goals. The role of portfolio management is to align the company's product offerings with the changing needs of its customers and the market while also ensuring that the portfolio delivers the desired financial performance and supports the company's overall business strategy [6-8]. Through portfolio management, companies can optimize resource utilization, mitigate risks, ensure alignment with business objectives, and facilitate resource allocation, resulting in more efficient and effective goal attainment [9].

The telecommunications industry has experienced significant changes in recent years due to privatization and liberalization, resulting in a dynamic and fast-growing environment. To remain competitive in this industry, companies require access to cutting-edge technology, innovations, and domestic and international market access. Effective portfolio management is essential for telecommunications companies to achieve their goals more efficiently and effectively, driving long-term success. Therefore, this research focuses on the telecommunications industry to investigate the challenges of portfolio management and identify successful practices that can enable companies to stay competitive and achieve their strategic goals.

This article presents a case study of a project portfolio management process (PPM) using the Stage-Gate model in a leading telecommunications company based in Serbia operating in a regional market. The study aims to contribute to the broader literature on portfolio management in the telecommunications industry by focusing on value capture and providing insights and lessons from the company's portfolio management process. Specifically, the article uses a unique Stage-Gate model to observe a company's organizational approach to achieving strategic goals and adapt project management according to the project's size, complexity, and characteristics.

This article makes a significant contribution to the existing literature on portfolio management. Its importance lies in its contribution to understanding the value of portfolio management in the telecommunications industry, particularly through the unique Stage-Gate PPM model. Companies in the industry can apply the insights and lessons learned from this case study to achieve their strategic goals and enhance their portfolio management processes.

The remainder of this article is structured as follows. The second section presents the literature review on PPM, rethinking value creation through PPM, the benefits of using the Stage-Gate model for PPM, and the fundamental PPM challenges in the telecommunication industry. The third section details the case study of introducing Stage-Gate PPM in a telecommunication company. The fourth section presents the case study's lessons learned. Finally, the paper concludes with a summary of the key points and provides directions for further research.

2. Literature review

2.1 Project Portfolio Management (PPM)

A portfolio represents a collection of projects, programs, subsidiary portfolios, and operations managed to achieve strategic objectives. Portfolio components (projects, programs, subsidiary portfolios, and operations) compete for a share of limited resources. Organizations must examine their unique circumstances and determine how to optimize and balance the portfolio components.

The beginnings of PPM date back to the 1950s in determining inventory portfolios [10]. PPM can be seen as one large entity of all the projects in an organization managed and sponsored by company managers. The right projects must be selected, prioritized and evaluated to form one entity that helps the organization reach its strategic goals. Project and program management are focused on process efficiency, but portfolio management refers to effectiveness – aligning tasks and priorities according to the strategic goals and vision of the company [10]. PPM attempts to answer the questions such as "What should we take on? What should be terminated? What is possible? What is needed?" [11].

Organizational strategy is composed of goals and policies that provide the overall direction and focus of the organization, as well as plans and actions to achieve those goals. PPM is the centralized management of one or more portfolios to achieve strategic objectives. Applying portfolio management principles aligns the portfolio and its components with the organizational strategy, contributing to the company's competitiveness and business success [12]. PPM can also be viewed as a dynamic activity through which an organization invests resources to achieve its strategic objectives by identifying, categorizing, monitoring, evaluating, integrating, selecting, prioritizing, optimizing, balancing, authorizing, transitioning, controlling, and terminating portfolio components. PPM is recognised as an active decision-making procedure that modifies a set of projects, whereby a business's list of active new projects is constantly updated and revised [10]. Optimising the PPM emphasises transparency through clear goals, roles and processes. Also, without standardization of project management, PPM is elusive [6].

2.2 Rethinking value creation through PPM: An organizational approach to achieving strategic objectives

One of the noteworthy changes in the project management field is the orientation of project management towards value delivery, defined as set of strategic activities contributing to achieving strategic goals and enhancing the company's operations. The concept of value capture has garnered considerable attention from scholars in the past decade [14]. However, a detailed examination of the value approach to project management has only recently emerged [15]. Value capture studies are pertinent because organizations have often encountered challenges capturing value from their projects [14]. These difficulties stem from an uncertain environment that necessitates anticipating project execution, managing uncertainties, and comprehending various stakeholders' diverse perspectives [16].

The shift from "product creation" to "value creation" has arisen from long-standing discussions about measuring project success [17]. The conventional observation of project success was centred on the well-known project management "Golden Triangle" imperatives – cost (remaining on a budget), time (meeting deadlines), and scope (meeting requirements). Rethinking project success assessment has spurred the recent emphasis on value and emphasized the importance of situating project management in a strategic context, namely emphasizing the importance of PPM [17]. PPM should ensure that the portfolio components (projects) that contribute the most to the value chain, impact the enhancement of the company's operations, and achieve strategic goals are selected for implementation. PPM provides a comprehensive framework for strategy execution and necessitates constant alignment with strategic objectives through processes and practices for project and program management to achieve strategic business objectives [7].

Effective PPM is critical for companies to achieve their strategic goals and deliver value to stakeholders beyond just commercial success [17] This requires careful consideration of both tangible and intangible benefits for stakeholders delivering the requisite values that stakeholders expect [18], and an alignment of strategy and execution while balancing feasibility with necessity. components that contribute most to a company's strategy. Although these components may be independent, they often compete for the same limited resource pools, making it crucial for companies to assess their resource potential and optimize the balance of their portfolio components. Once the projects are selected, it is imperative to allocate limited resources to them in a coordinated manner to achieve the best possible results [7, 19]. Coordinated management makes it possible to allocate human, financial and physical (logistical/material) to achieve the best possible results. In other words, effective PPM ensures that a company's strategic goals are met by selecting, allocating resources to, and implementing the most important projects.

By analysing portfolio components and their interdependencies, PPM can identify potential challenges and risks, constantly updating and revising an active list of projects [20, 21]. This helps make strategic decisions that align with the company's goals. To achieve this, PPM must be integrated with organizational planning and business analysis to analyse current business risks, which may drive strategic changes to support planned portfolio components.

Portfolio components are grouped based on risk, financing, and other parameters to facilitate effective work management while achieving organizational strategies and priorities. Assigning appropriate priorities to portfolio components is critical to managing the portfolio effectively. PPM enables a company to execute the right projects at the right time by selecting and managing projects as a portfolio of investments. Linking PPM to strategy balances resource utilization and investments to maximize the value delivered in executing programs, projects, and operational activities. PPM has become a key element of how companies deliver value to multiple stakeholders and achieve business success.

However, achieving strategic goals and delivering value through PPM are not without their challenges. These include obtaining and maintaining senior management support, determining short-term and long-term goals, managing limited resources and capabilities, and sustaining the ability to execute [17]. Hence, companies must focus on doing the right projects at the right time in the right way to avoid the wastage of precious resources [22].

2.3 Benefits of using the Stage-Gate model for PPM

The Stage-Gate model, which was introduced in the mid-1980s, has helped many companies making the new product development process more effective and built to prevent or minimise risks [23]. The model has been continually refined and enhanced by industry leaders to make it more flexible, scalable, and adaptable, incorporating best practices such as better governance, integration with portfolio management, accountability, and continuous improvement.

The Stage-Gate model's benefits extend to portfolio management, where it provides a systematic approach to managing projects and ensuring alignment with a company's strategic objectives while achieving the desired outcomes. Adopting a systems perspective to the Stage-Gate process underscores the importance of feedforward controls, such as planning and forecasting, in making critical decisions such as go/kill/hold/recycle [24, 25].

This article highlights the benefits of using the Stage-Gate model for portfolio management, emphasizing its ability to improve decision-making, enhance resource allocation, mitigate risks, and improve communication. By using predefined criteria and gates, decision-makers can objectively assess the project's potential and determine whether to continue, revise, or terminate it. The strength of Stage-Gate methodology is its simplicity and decision-making based on information available at that moment [26].

The Stage-Gate model enables effective resource allocation, ensuring each project has the resources it needs to succeed. This is achieved by using predefined stages and gates that help companies allocate resources based on the project's needs and potential. For integrating the PPM to the Stage-Gate process, gates have to be modified. In addition, resource allocation methods are added to the gates without reprioritising the entire set of projects every month.

The model also includes a structured approach to risk management, enabling companies to identify and address potential risks at each gate, thereby reducing the likelihood of project failure.

Effective communication and reporting are also integral parts of the Stage-Gate model, ensuring stakeholders are informed and engaged throughout the project's lifecycle. This helps to build trust and support for the project, increasing the chances of success.

2.4 PPM in the telecommunication industry

The telecommunication industry is highly competitive and subject to rapid technological changes. Consequently, telecommunication companies invest substantial amounts of capital into numerous projects, such as network infrastructure, product development, and marketing initiatives.

Effective PPM within this industry encompasses managing the company's portfolio of products and services, including developing and launching new products, optimizing existing products, and discontinuing underperforming ones. The primary objective of PPM is to identify and prioritize projects that are most likely to yield the highest return on investment, ensuring that the company's resources are utilized efficiently and effectively to maximize value. PPM optimize a portfolio's value, develop its strategic alignment, and balances its assignments [15]. In addition, it helps to diversify the company's portfolio of projects, reducing the risk of a single project failure affecting the entire business. It also enables the company to monitor the progress of each project and make adjustments if necessary, reducing the likelihood of unexpected setbacks.

In the telecommunications industry, effective PPM must ensure that the company's product offerings align with the dynamic needs of the customers and the market while delivering the desired financial performance and supporting the company's overall business strategy [27]. This ensures that all projects are working in unison towards a common objective, thus enhancing the likelihood of success. Portfolio management is instrumental in efficiently allocating company resources, such as capital, staff, and time, across different projects. This not only ensures that each project has the resources it needs to succeed but also prevents resource bottlenecks and wastage.

Peter Krüssel [28] has identified various challenges faced by telecommunication companies in managing their portfolio of products and services. One major challenge is the rapidly changing technology landscape, which requires companies to adjust their portfolio offerings to stay relevant continuously [29]. Moreover, telco companies require significant capital investments to build and maintain their networks, making investing in new products or services that may not have a guaranteed return on investment difficult. In addition, increasing competition from new entrants such as over-the-top (OTT) providers and technology companies also poses challenges for maintaining a competitive portfolio and differentiating themselves from their competitors. Another challenge is balancing short-term financial goals with longer-term strategic objectives. This can create tension between investing in new products and services that may have a longer-term payoff versus maintaining profitability in the short term. Furthermore, the highly regulated nature of the telecommunication industry with complex regulations that vary by country and region presents challenges in introducing new products or services. As a result, companies need to navigate regulatory requirements and obtain necessary approvals, which can be time-consuming and expensive.

The ultimate goal of linking PPM with organisational strategy and strategic business execution is to establish a balanced, realistic plan that will help achieve both short-term financial goals and long-term strategic goals [30] .Furthermore, successful companies in this industry must be agile and innovative [31, 32], which is unthinkable without a well-established PPM.

3. Case study: Introducing Stage-Gate PPM in telecommunication company

The case study presented in this section focuses on successful implementation of a PPM process using the Stage-Gate model in a prominent telecommunications company based in Serbia. The company offers a range of services, including mobile and fixed-line telephony, broadband internet, and cable television, and has a strong market position with over 11 million subscribers in three regional markets.

3.1 Introducing PPM in the portfolio cycle

Like programs and projects, portfolios require diligent life cycle management, including initiation, planning, execution, and optimization, to ensure stability and adaptability in a constantly changing environment. The portfolio's performance is monitored and controlled during all phases, with relevant information analysed and decisions made about which components should be transferred to the next stage. The phases in the portfolio life cycle are not necessarily sequential and can have multiple iterations in one phase, and all phases within the portfolio are subject to change. Management decisions are made within the portfolio's life cycle, enabling the portfolio to be changed and updated to adapt to internal and external factors. For instance, if there are legal and regulatory changes, the PPM process must align the management plans with the new requirements, as illustrated in Fig. 1.

The goal of establishing the PPM process within this telecommunication company under analysis was to identify all business requirements at the company level and enable centralized management of priorities for programs, projects, and other business-as-usual (BaU) requirements derived from the current business strategy and market demands. A PPM process was deemed crucial to prepare the organization for future exploration of new markets and development possibilities, acquiring new skills and competencies, and adapting to rapidly changing business environments [8].

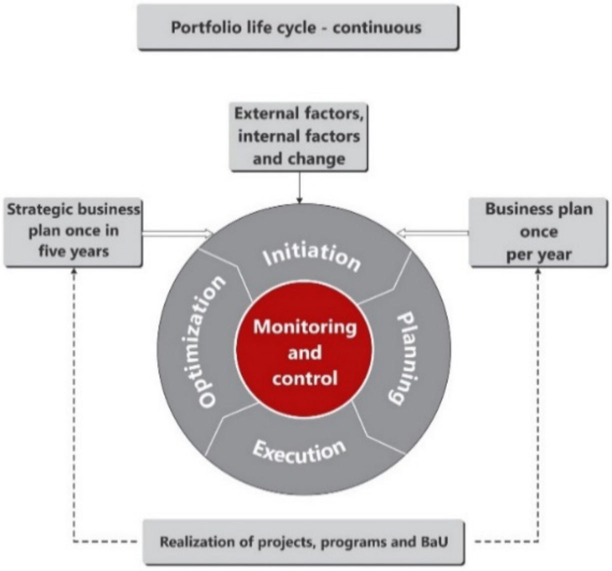


Fig. 1 Portfolio Life Cycle as a continuous process

The process aimed to shift the organization's mindset from tactical to strategic by selecting and managing projects aligned with its strategy. The company's business strategy, formulated by the top management, clearly indicated how the company's operations would be shaped in the following period. PPM was central to achieving the organization's intended strategies and future growth [11].

In 2022, the PPM process was introduced through a pilot project, which was based on the five-year strategic business plan that had been created in 2021. The pilot project covered several operational fields the Portfolio Governance Board had to deal with, including gaining an end-to-end overview of all inserted portfolio components and their associated activities. The pilot project also involved prioritizing the realization of portfolio components/activities and understanding the possible consequences of not completing other activities due to priority changes. Additionally, the pilot project aimed to balance short-term and long-term portfolio components/activities, present strategically important projects and programs and connect the portfolio components with monitoring the implementation of the strategic business plan. The monitoring involves using defined strategic initiatives, key performance indicators (KPIs), and preventive/corrective measures. Doing so makes it possible to properly balance and prioritize projects/programs/BaU within the established portfolio, which in turn contributes to achieving the defined strategic goals. It is essential to base the KPIs and strategic initiatives on the company's strategic goals and ensure that they contribute to its overall success and sustainability [31].

3.2 Stage-Gate project Portfolio Management model

In this case study, the Stage-Gate model is used to establish the PPM process by taking a big-picture overview of all activities at the company level, dividing their realization into stages, and enabling an easier decision-making process in every gate (determine whether or not the project can continue to the next phase). The Stage-Gate PPM model presented in this article is uniquely valuable because it can be adapted to different types of projects [33]. It shares some similarities with the Agile-Stage-Gate model, as it incorporates elements of Agile approaches such as Sprint meetings and artefacts [34].

The Stage-Gate PPM model description

I) STAGE 1 (S1): Proposal Selection
Telecommunication company receives a large number of proposals from its different organizational units for potential projects and programs. These proposals cover various areas such as network expansion, new product development, and process improvement initiatives. To start the PPM process, the Executive Directors of each organizational unit

consider nominated proposals in their respective units of responsibility and nominate the most promising ones for further selection at the company level. The output of the S1 is a List of company-level nominated demands for projects/programs (Gate 1 – G1).

II) STAGE 2 (S2): Selection of Nominated Demands
The company conducts a business analysis and financial asset verification for the nominated demands to determine which demands should be realized as projects or programs. The output of the S2 is a List of portfolio components at the company level, which are entered as portfolio components in the Portfolio Database (Gate 2 – G2).

III) STAGE 3 (S3) Prioritization
Prioritization is carried out iteratively during Portfolio Sprint Meetings. The Portfolio Manager and Portfolio Team prepare relevant information and proposals to aid decision-making. The Portfolio Governance Board reviews the proposals and makes decisions on prioritization, ensuring that Go-NoGo decisions are clearly defined when priorities change. Executive Directors are also involved in decision-making for reprioritization or termination of activities when necessary. The goal of prioritization is to ensure that the implementation of short-term and long-term activities is coordinated to bring the greatest value for the company, taking into account the proper allocation of time, money and human resources to individual activities. A detailed Risk/Value analysis is conducted to confirm project priorities through discussion at Portfolio Sprint meetings. This process aims to ensure that the company is realizing the most suitable projects by carefully calibrating, levelling, and harmonizing R/V parameters at the company level through a joint and transparent assessment. The output of the S3 is a List of prioritized portfolio components at the company level (Gate 3 – G3).

IV) STAGE 4 (S4): Categorization of the Projects
The Program Management Team utilizes the List of prioritized portfolio components (G3) to categorize the projects and recommend an appropriate project methodology for each. The team considers various factors, including project scope, complexity, and duration. Still, the most important is the benefit that the project brings to the company in the short and long term to classify the projects into small, medium, or large categories. This categorization allowed for a balanced distribution of short-term and long-term activities as part of the PPM process.

Fig. 2 present the modified Stage-Gate model integrating the PPM.

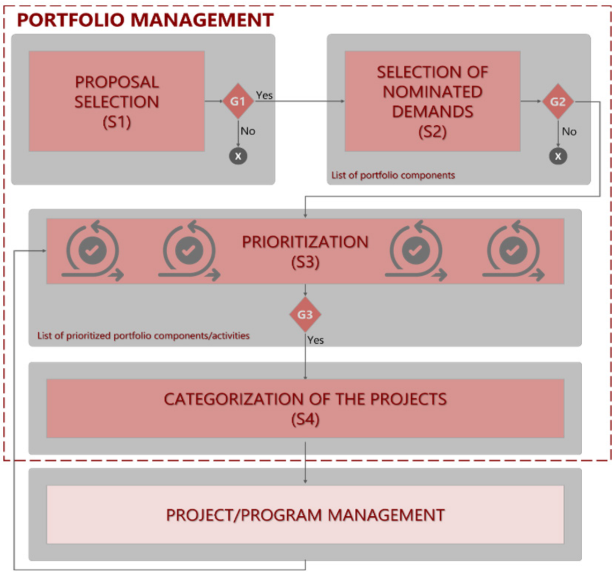


Fig. 2 Stage-Gate PPM Model

For each project category, the Program Management Team analyse each project individually and proposed an appropriate project methodology. They considered the need for iterative and incremental development throughout the project life cycle, changing customer requirements, uncertain processes, and value delivery. As no one-size-fits-all project methodology suits all companies and project types, the team recommended either a waterfall approach, an agile approach, or a hybrid methodology that incorporated both waterfall and agile approaches. Agile approaches are better suited for situations where changing customer requirements and uncertain processes are expected, and value delivery is a focus. On the other hand, traditional project management approaches, such as the Waterfall methodology, are plan-driven, focused on detailed planning and process control, and customer requirements are defined at the outset [35].

3.3 Stage-Gate PPM model roles and responsibilities

The Stage-Gate PPM process involves several key participants, including Executive Directors, the Portfolio Governance Board, Middle Management – other department directors who are not members of the Portfolio Governance Board but whose participation is relevant in portfolio management, and a Permanent Team from the Strategy Department that supports the work of the Portfolio Governance Board.

1. *The Executive Directors* are responsible for managing the participation of the company's organizational units in the PPM process. They make decisions on which requests will be nominated to achieve strategic goals based on business needs and potential, and also monitor and control the realization of portfolio components. They may also decide on reprioritization or termination of portfolio components if necessary.
2. *The Portfolio Governance Board* is composed of 10 department directors from organizational units relevant to portfolio management, who have the right to vote for prioritization/reprioritization of activities (projects/programs/BaU). These departments include product development, technical and IT, sales, customer care, financial controlling, and budgeting. The Portfolio Governance Board is coordinated by the director of the Strategy department, who also oversees periodic Portfolio Sprint Meetings.
3. *Middle Management* comprises other department directors who are not members of the Portfolio Governance Board but participate in all or some activities based on their competencies and needs. They do not have the right to vote for prioritization/reprioritization but are responsible for communicating with their organizational units about any changes in prioritization status.
4. *The Permanent Team* from the Strategy Department supports the work of the Portfolio Governance Board and the preparation of the Portfolio Sprints. It includes the Portfolio Manager, Portfolio Team, Strategic Team, and Program Management Team.
   - The Portfolio Manager and Portfolio Team analyse all relevant information and prepare a proposal for making decisions on reprioritization of portfolio components at the next Portfolio Governance Board sprint. They also prepare a unified report on the status of the portfolio in Power BI, containing relevant indicators on the status and trend of realization on the project/program, and a proposal for making decisions on balancing and reprioritization.
   - The Strategic Team connects strategic initiatives with projects/programs/BaU entered in the Portfolio table and monitors the implementation of the current strategic business plan by quarter, strategic and operational KPIs, and the connection with the implementation of portfolio components.
   - The Program Management Team, in cooperation with Business requestors, is responsible for categorizing projects and selecting an adequate project management methodology.

Through open and transparent governance, including processes for categorizing, prioritizing, selecting, and approving portfolio components, PPM key stakeholders are more likely to accept the decisions and agree with the process, even when they may not fully endorse the decisions made.

3.4 Stage-Gate PPM model artefacts and meetings

The *Portfolio Database*, as a crucial PPM artefact, was established at the inception of the pilot project in Q1 2022 to record all company planned, ongoing or completed activities, including Program, Project, and BaU, along with relevant data. All responsible participants, including Portfolio Governance Board members and their associates, Project/Program managers/Product owners, Process managers, and the Strategic team, constantly update the Portfolio Database. Key data for each activity is arranged in columns, including the name of the aggregated activity, its type (PR/PG /BaU), description, owner, project manager/product owner, percentage of realization, status, strategic pillar/strategic initiatives, traffic light status (project health check), category, and risk/value. These data are essential for top management reporting and must be updated promptly to ensure accurate and complete reports.

Each portfolio component must be assigned to either the short-term category, which includes small application changes and small and medium projects, or the long-term category, which includes large projects and programs and BaU activities. The primary criterion for project categorization is the benefit it brings to the company in the short and long term.

All activities submitted by business requestors for prioritization using various tools must be added to the Portfolio Database. Business requestors can indicate the need to prioritize an activity by adding the "P" mark in the Portfolio Database and a brief explanation.

This database serves as a source for systematizing, monitoring, and reporting, as well as a basis for balancing and prioritizing activities during periodic meetings of the Portfolio Governance Board, known as Portfolio Sprints. This is a living artefact being continuously maintained.

The *Portfolio Sprint* is a regularly scheduled, time-limited meeting lasting one hour on Wednesdays. During this meeting, the Project Governance Body directs its attention towards pre-determined tasks focusing on achieving specific objectives.

The Portfolio Manager, along with the Portfolio Team and operational associates of department directors who serve as Portfolio Governance Board members, evaluate the impact of increasing the priority of specific portfolio components. They analyse all relevant information and prepare a proposal for making decisions on the reprioritization of portfolio components at the next Portfolio Governance Board sprint once a week. Based on their proposed new ranking, the final priority decision is made during the Portfolio Governance Board's Portfolio Sprint Meeting.

The Portfolio Sprint agenda often includes an end-to-end (E2E) overview of individual activities, which helps shift the focus from individual activities to the overall activity (Strategic business plan), making the realization process more comprehensive. To provide an E2E overview of portfolio components, it is necessary to aggregate all basic activities by linking them to the overall activity, enabling the monitoring of interdependence during the realization of associated activities.

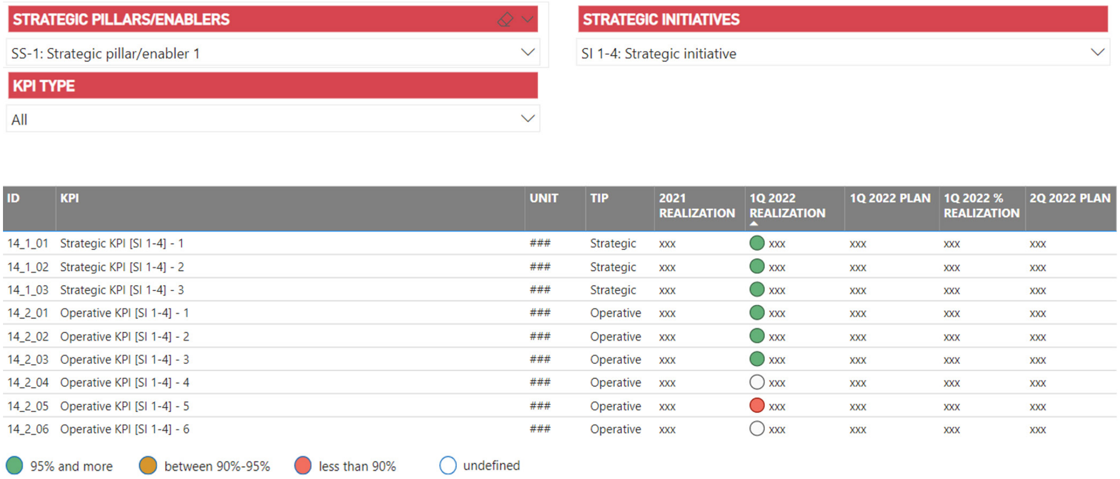
In addition, *Regular Status meetings* with the strategic initiatives' owners provide an opportunity to identify projects and programs that have strategic importance and should be prioritized.

Long-term activities identified through this process are specially flagged in the Portfolio Database to ensure they receive the necessary attention and resources. Meanwhile, strategic initiatives owners are also able to apply for and launch projects and programs that have not been recognized by the previous analysis but which they consider to be strategically important. This flexibility allows the continuous refinement of the strategic plan to adapt to changing circumstances.

This approach was beneficial in launching new services, as it allows monitoring the implementation of activities required for Customer Care, Business Analytics, and Marketing sectors after go-live. It also facilitates dividing complex, transformational programs into smaller, more manageable projects.

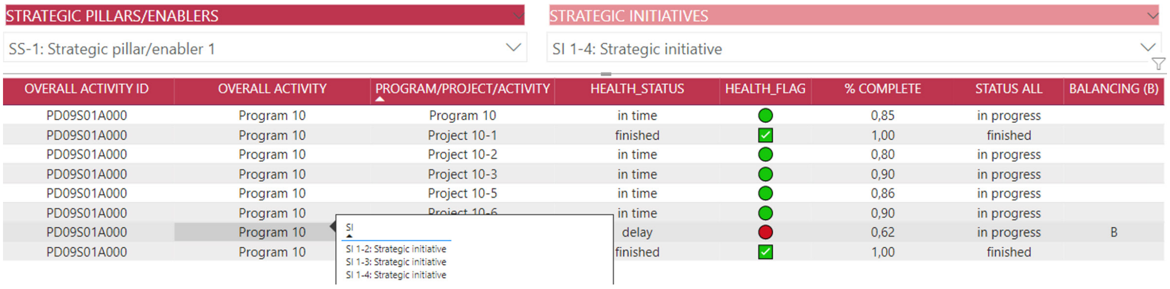
The Strategic team plays a critical role in providing ongoing support to the strategic initiative owner, helping them to monitor and report on the progress of their initiatives effectively. This coordination leads to the preparation of quarterly reports on the implementation of the strategic business plan, which serve as the basis for prioritizing projects and programs within the PPM process. The use of Power BI reports, as shown in Figures 3 and 4, further enhances the PPM process.

Fig. 3 shows a report for one strategic initiative (SI 1.4.) by portfolio components (6 projects are realizing according to plan, 1 project is late, and 2 projects have yet to start). Fig. 4 shows that the late project has the flag "B" in the Portfolio Database, indicating the need for balancing this project at the Portfolio Sprint meetings. In this way, the necessary feedback from monitoring the strategic plan's implementation is provided to the PPM process. This loop enables comprehensive calibration between portfolio components and strategy implementation results at the company level through a joint and transparent assessment, which contributes to realizing the right things in the right way.



ID	KPI	UNIT	TIP	2021 REALIZATION	1Q 2022 REALIZATION	1Q 2022 PLAN	1Q 2022 % REALIZATION	2Q 2022 PLAN
14_1_01	Strategic KPI [SI 1-4] - 1	###	Strategic	xxx	xxx	xxx	xxx	xxx
14_1_02	Strategic KPI [SI 1-4] - 2	###	Strategic	xxx	xxx	xxx	xxx	xxx
14_1_03	Strategic KPI [SI 1-4] - 3	###	Strategic	xxx	xxx	xxx	xxx	xxx
14_2_01	Operative KPI [SI 1-4] - 1	###	Operative	xxx	xxx	xxx	xxx	xxx
14_2_02	Operative KPI [SI 1-4] - 2	###	Operative	xxx	xxx	xxx	xxx	xxx
14_2_03	Operative KPI [SI 1-4] - 3	###	Operative	xxx	xxx	xxx	xxx	xxx
14_2_04	Operative KPI [SI 1-4] - 4	###	Operative	xxx	xxx	xxx	xxx	xxx
14_2_05	Operative KPI [SI 1-4] - 5	###	Operative	xxx	xxx	xxx	xxx	xxx
14_2_06	Operative KPI [SI 1-4] - 6	###	Operative	xxx	xxx	xxx	xxx	xxx

Fig. 3 Example of Strategic business plan report (KPIs)



OVERALL ACTIVITY ID	OVERALL ACTIVITY	PROGRAM/PROJECT/ACTIVITY	HEALTH_STATUS	HEALTH_FLAG	% COMPLETE	STATUS ALL	BALANCING (B)
PD09S01A000	Program 10	Program 10	in time		0,85	in progress	
PD09S01A000	Program 10	Project 10-1	finished		1,00	finished	
PD09S01A000	Program 10	Project 10-2	in time		0,80	in progress	
PD09S01A000	Program 10	Project 10-3	in time		0,90	in progress	
PD09S01A000	Program 10	Project 10-5	in time		0,86	in progress	
PD09S01A000	Program 10	Project 10-6	in time		0,90	in progress	
PD09S01A000	Program 10		delay		0,62	in progress	B
PD09S01A000	Program 10		finished		1,00	finished	

Fig. 4 Example of Strategic business plan report (Overall Activity)

Through this continuous process of analysis and refinement, PPM enables the company's operations to remain flexible and responsive to changing circumstances, thereby increasing the likelihood of successfully implementing the desired strategy.

3.5 Balancing long-term and short-term activities in the Stage-Gate PPM model

Balancing long-term and short-term activities involves carefully considering risk and value to prioritize projects based on their potential impact on the company [36]. This process aims to ensure that the company realizes the most suitable projects by carefully calibrating levelling [17], and harmonizing Risk/Value parameters at the company level through a joint and transparent assessment.

To achieve this, a Risk/Value model has been introduced, which involves assessing each project based on two criteria: value criteria for calculating the value of portfolio components which are determined by the company's strategic plan, and risk criteria for risk calculation which are determined based on architectural, organizational and investment complexity.

By calculating each project's value and risk scores and creating a Risk/Value Bubble diagram, it is possible to identify which projects are of the highest priority. The diagram is divided into four quadrants: (1) generic – low value and low risk, (2) bottleneck – low value and high risk, (3) leverage – high value and low risk, and (4) critical – the performance and value of the services in customers' hands; high value and high risk [34]. The bubble size corresponds to the project size (small, medium, large), and its position on the diagram reflects the calculated Risk/Value ratio. In Fig. 5, the example of the Risk/Value Bubble diagram (Scatter chart in Power BI) is given.

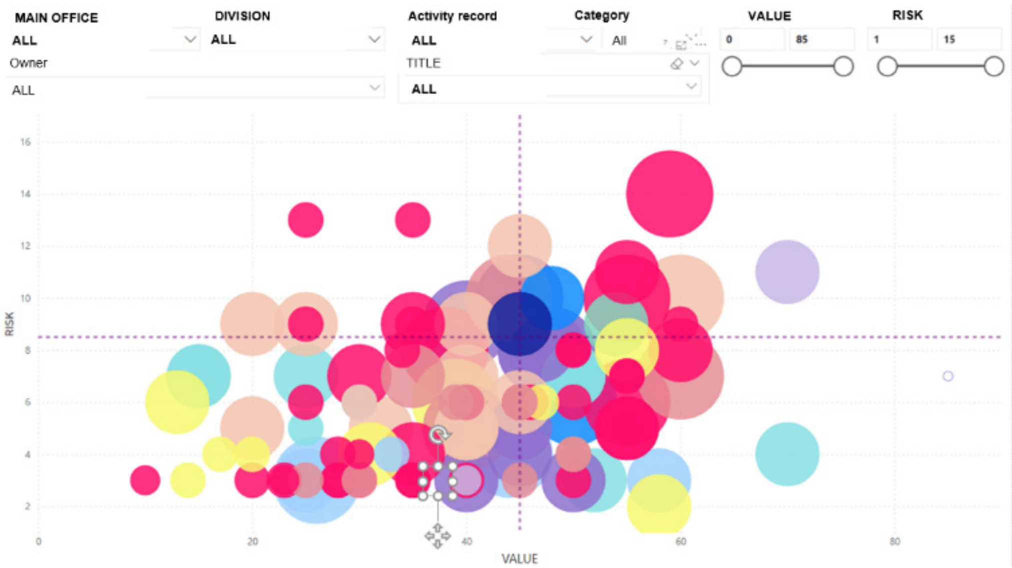


Fig. 5 Risk/Value diagram – Bubble diagram (Scatter chart in Power BI)

To achieve a balanced portfolio, it is essential to allow for the temporary postponement of long-term activities to prioritize short-term activities. During the prioritization process, evaluating the impact of any such postponement on the initial deadline of the long-term activity and maintaining a record of such changes is necessary. The aspiration is to achieve maximum portfolio components with a high-value and low-risk profile. This is often achieved with limited financial, human, and other resources, benefiting the company significantly. Activities with lower value should not be pursued if they entail a high level of risk. In such cases, it is imperative to minimize the risk by simplifying the requests or abandoning their realization altogether.

4. Key lessons learned

Key lessons learned from implementing Stage-Gate PPM are as follows:

- Top management support is essential for successfully establishing PPM process, but a bottom-up approach is necessary to include all company activities as portfolio components. It is critical to involve all relevant stakeholders in the process, as they are the ones who will ultimately contribute to the portfolio. A bottom-up approach ensures that all activities and projects are considered in the portfolio management process, which can lead to a more comprehensive and effective portfolio.
- It is necessary to define the Portfolio Governance Board to govern portfolio realization and provide centralized value and risk management, as well as optimize resources between portfolio components, with the support of portfolio managers and the portfolio team. The Portfolio Governance Board plays a critical role in PPM, providing oversight and guidance on portfolio components.
- The proposed PPM Stage-Gate model used in this case study is unique because it incorporates elements of Agile approaches such as Sprint meetings and artifacts and it can be adapted to different types of projects. The model is suitable for enabling iterative decision-making by narrowing choices in consequential phases. By breaking down the process into sequential stages, the PPM Stage-Gate model allows for a more focused and structured approach to decision-making, ensuring that each portfolio component is thoroughly evaluated before moving to the next stage. Through the continuous refinement in iterations in the third phase Stage-Gate PPM enables the company's operations to remain flexible and responsive to changing circumstances, thereby increasing the likelihood of successfully implementing the desired strategy. Additionally, the model is fully customizable in its fourth phase, allowing for flexibility and tailored solutions to each project's needs.

- A prerequisite for well-defined criteria of the Risk/Value model is a clearly defined strategic framework because the assessments of risks and values are based on the elements of the strategic framework (values are related to strategic initiatives and risks to complexity). The Risk/Value model is a widely used approach for evaluating portfolio components, but its effectiveness depends on the quality of the criteria used to assess risk and value.
- To ensure adequate scoring regarding the value of a portfolio component, it is necessary that the criteria for the valuation are tied to different strategic objectives (mutually exclusive). For example: prohibiting the same project improves revenue and user experience simultaneously. By ensuring that the criteria used for evaluation are mutually exclusive and tied to different strategic objectives, organizations can prioritize portfolio components that are most relevant to their goals.
- To monitor the realization of strategic goals, it is necessary to carry out balancing projects and programs based on the achieved values of KPIs of the strategic initiatives. In contrast, feedback for balancing short-term and long-term portfolio components is obtained from monitoring the implementation of the strategic plan. By using KPIs to evaluate the achievement of strategic initiatives, organizations can identify which projects and programs are contributing to the realization of strategic goals. By monitoring the implementation of the strategic plan, organizations can ensure that short-term and long-term portfolio components are appropriately balanced.
- A centralized data collection and processing system for reporting on portfolio components must be provided. Maintenance of the Portfolio Database is essential for correct decision-making. By ensuring that data is collected and processed centrally, organizations can make more informed decisions about portfolio components and their alignment with organizational objectives. By maintaining the portfolio database, organizations can ensure that portfolio components are evaluated consistently over time.

5. Conclusion and further research

The telecommunications industry faces various challenges, and successful companies must be innovative, agile and know how to operate with data. The Stage-Gate PPM model can help develop a new product/service or process improvement, but companies must be creative and know how to adapt it to their needs. Using E2E overview and unification at the level of aggregated activities, as well as linking portfolio components with monitoring the implementation of strategic initiatives within the strategic business plan, enabled the initial establishment of the portfolio balancing process.

The case study presented in this article highlights the importance of top management support, a well-defined governance structure, the use of the Risk/Value model, monitoring of KPIs, and a centralized data collection and processing system for successful Stage-Gate PPM.

This study contributes to the growing body of literature on PPM and provides practical insights into the integration of PPM with the Stage-Gate model. Practical examples of PPM integration with the Stage-Gate model are the missing link in the academic community researching PPM as a critical component of companies' success based on value delivery.

However, this study has some limitations. The case study was conducted in a single telecommunication company, and the findings may not be generalizable to other industries or companies. Further research is needed to test the effectiveness of the proposed Stage/Gate PPM model integration in other industries and companies. Additionally, the study focused on the initial implementation of PPM, and further research is needed to examine the long-term effects of PPM on business performance. Finally, this study did not assess the financial implications of implementing PPM, and future research should examine the impact of PPM on organizational performance.

Finally, future research should also focus on improving the Stage-Gate PPM model with methodologies, tools, and practices for building a sustainable business. For example, the Lean start-up and Design thinking methodologies could be integrated into the first stage of the Stage-Gate PPM model to enhance the ideation and validation of new ideas. Using AI in the second stage could help improve decision-making and reduce time and resource consumption. Finally, ESG factors and project resilience could be incorporated into the third and fourth stages to ensure the sustainable development of projects and the company as a whole.

Acknowledgement

This research has been supported by the Ministry of Science, Technological Development and Innovation through project no. 451-03-47/2023, 01/200156 "Innovative scientific and artistic research from the FTS (activity) domain".

References

[1] Grand view research. Telecom services market size, share & trends analysis report by service type (mobile data services, machine-to-machine services), by transmission (wireline, wireless), by end-use, by region, and segment forecasts, 2023 – 2030, from *https://www.grandviewresearch.com/industry-analysis/global-telecom-services-market*, accessed February 15, 2023.

[2] Salunkhe, O., Berglund, Å.F. (2022). Industry 4.0 enabling technologies for increasing operational flexibility in final assembly, *International Journal of Industrial Engineering and Management*, Vol. 13, No. 1, 38-48, doi: 10.24867/IJIEM-2022-1-299.

[3] Ciric, D., Lolic, T., Gracanin, D., Stefanovic, D., Lalic, B. (2020). The application of ICT solutions in manufacturing companies in Serbia, In: Lalic, B., Majstorovic, V., Marjanovic, U., von Cieminski, G., Romero, D. (eds.), *Advances in production management systems, Towards smart and digital manufacturing, APMS 2020, IFIP advances in information and communication technology*, Vol. 592, Springer, Cham, Switzerland, 122-129, doi: 10.1007/978-3-030-57997-5_15.

[4] Fiedler, S. (2010). Managing resistance in an organizational transformation: A case study from a mobile operator company, *International Journal of Project Management*, Vol. 28, No. 4, 370-383, doi: 10.1016/j.ijproman.2010.02.004.

[5] Nagy, G., Vida, G., Boros, L., Ciric, D. (2019). Decision trees in environmental justice research – A case study on the floods of 2001 and 2010 in Hungary, *Open Geosciences*, Vol. 11, No. 1, 1025-1034, doi: 10.1515/geo-2019-0079.

[6] Teller, J., Unger, B.N., Kock, A., Gemünden, H.G. (2012). Formalization of project portfolio management: The moderating role of project portfolio complexity, *International Journal of Project Management*, Vol. 30, No. 5, 596-607, doi: 10.1016/j.ijproman.2012.01.020.

[7] Petit, Y. (2012). Project portfolios in dynamic environments: Organizing for uncertainty, *International Journal of Project Management*, Vol. 30, No. 5, 539-553, doi: 10.1016/j.ijproman.2011.11.007.

[8] Rank, J., Unger, B.N., Gemünden, H.G. (2015). Preparedness for the future in project portfolio management: The roles of proactiveness, riskiness and willingness to cannibalize, *International Journal of Project Management*, Vol. 33, No. 8, 1730-1743, doi: 10.1016/j.ijproman.2015.08.002.

[9] Jerbrant, A., Karrbom Gustavsson, T. (2013). Managing project portfolios: Balancing flexibility and structure by improvising, *International Journal of Managing Projects in Business*, Vol. 6, No. 1, 152-172, doi: 10.1108/17538371311291071.

[10] Danesh, D., Ryan, M.J., Abbasi, A. (2018). Multi-criteria decision-making methods for project portfolio management: A literature review, *International Journal of Management and Decision Making,* Vol. 17, No. 1, 75-94, doi: 10.1504/IJMDM.2018.088813.

[11] Lappi, T.M., Aaltonen, K., Kujala, J. (2019). Project governance and portfolio management in government digitalization, *Transforming Government: People, Process and Policy*, Vol. 13, No. 2, 159-196, doi: 10.1108/TG-11-2018-0068.

[12] Silva, C.S., Pereira, C., Magano, J. (2023). The value of project management to competitiveness: Key factors from a holistic and practical perspective, *International Journal of Managing Projects in Business*, Vol. 16, No. 1, 67-91, doi: 10.1108/IJMPB-02-2020-0042.

[13] Chang, A., Chih, Y.-Y., Chew, E., Pisarski, A. (2013). Reconceptualising mega project success in Australian defence: Recognising the importance of value co-creation, *International Journal of Project Management*, Vol. 31, No. 8, 1139-1153, doi: 10.1016/j.ijproman.2012.12.005.

[14] Uzakova, S.T., Shildibekov, Y.Z. (2021). Value in project management: New direction in theory and practice, *Central Asian Economic Review,* No. 6, 116-124, doi: 10.52821/2789-4401-2021-6-116-124.

[15] Bos-de Vos, M., Volker, L., Wamelink, H. (2019). Enhancing value capture by managing risks of value slippage in and across projects, *International Journal of Project Management*, Vol. 37, No. 5, 767-783, doi: 10.1016/j.ijproman.2018.12.007.

[16] Green, S.D., Sergeeva, N. (2019). Value creation in projects: Towards a narrative perspective, *International Journal of Project Management*, Vol. 37, No. 5, 636-651, doi: 10.1016/j.ijproman.2018.12.004.

[17] Maslennikov, V.V., Popova, E.V., Kalinina, I.A. (2022). Classic project management based on PMBOK 7.0, In: Popkova, E.G., Polukhin, A.A., Ragulina, J.V. (eds.), *Towards an increased security: Green innovations, intellectual property protection and information security*, *Lecture notes in networks and systems,* Vol. 372, 835-840, Springer, Cham, Switzerland, doi: 10.1007/978-3-030-93155-1_90.

[18] Martinsuo, M. (2013). Project portfolio management in practice and in context, *International Journal of Project Management*, Vol. 31, No. 6, 794-803, doi: 10.1016/j.ijproman.2012.10.013.
[19] Müller, R., Martinsuo, M., Blomquist, T. (2008). Project portfolio control and portfolio management performance in different contexts, *Project Management Journal,* Vol. 39, No. 3, 28-42, doi: 10.1002/pmj.20053.
[20] Gutiérrez, E., Magnusson, M. (2014). Dealing with legitimacy: A key challenge for Project Portfolio Management decision makers, *International Journal of Project Management*, Vol. 32, No. 1, 30-39, doi: 10.1016/j.ijproman.2013.01.002.
[21] Pinto, G., Silva, F.J.G., Fernandes, N.O., Casais, R., Baptista, A., Carvalho, C. (2020). Implementing a maintenance strategic plan using TPM methodology, *International Journal of Industrial Engineering and Management*, Vol. 11, No. 3, 192-204, doi: 10.24867/IJIEM-2020-3-264.
[22] Aristodemou, L., Tietze, F., Shaw, M. (2020). Stage gate decision making: A scoping review of technology strategic selection criteria for early-stage projects, *IEEE Engineering Management Review*, Vol. 48, No. 2, 118-135, doi: 10.1109/EMR.2020.2985040.
[23] Cooper, R.G. (1990). Stage-gate systems: A new tool for managing new products, *Business Horizons*, Vol. 33, No. 3, 44-54, doi: 10.1016/0007-6813(90)90040-I.
[24] Baker, M., Bourne, M. (2014). A governance framework for the idea-to-launch process: Development and application of a governance framework for new product development, *Research-Technology Management*, Vol. 57, No. 1, 42-48, doi: 10.5437/08956308X5701105.
[25] Chenger, D., Woiceshyn, J. (2021). Executives' decision processes at the front end of major projects: The role of context and experience in value creation, *Project Management Journal*, Vol. 52, No. 2, 176-191, doi: 10.1177/8756972820977225.
[26] Weber, Y., Tarba, S.Y. (2014). Strategic agility: A state of the art, Introduction to the special section on strategic agility, *California Management Review*, Vol. 56, No. 3, 5-12, doi: 10.1525/cmr.2014.56.3.5.
[27] Krüssel, P. (2019). *Future Telco, Successful positioning of network operators in the digital age,* Springer, Cham, Switzerland, doi: 10.1007/978-3-319-77724-5.
[28] Kumar, R., Bose, P. (2022). Case study Telecom industry and competitive landscape in India: Will MTNL and BSNL successfully recover?, *IIM Ranchi Journal of Management Studies,* Vol. 1, No. 1, 82-98, doi: 10.1108/irjms-12-2021-0179.
[29] Mustonen, E., Seppänen, J., Tolonen, A., Harkonen, J., Haapasalo, H. (2020). Product portfolio management strategic targets and kpis over life-cycle: A case study in telecommunications business, *Managing Global Transitions,* Vol. 18, No. 1, 5-23, doi: 10.26493/1854-6935.18.5-23.
[30] Basulo-Ribeiro, J., Amorim, M., Teixeira, L. (2023). How to accelerate digital transformation in companies with Lean Philosophy? Contributions based on a practical case, *International Journal of Industrial Engineering and Management*, Vol. 14, No. 2, 94-104, doi: 10.24867/IJIEM-2023-2-326.
[31] Kim, M., Chai, S. (2022). The role of agility in responding to uncertainty: A cognitive perspective, *Advances in Production Engineering And Management*, Vol. 17, No. 1, 57-74, doi: 10.14743/apem2022.1.421.
[32] Cooper, R.G., Sommer, A.F. (2016). The agile – stage-gate hybrid model: A promising new approach and a new research opportunity, *Journal of Product Innovation Management*, Vol. 33, No. 5, 513-526, doi: 10.1111/jpim.12314.
[33] Brock, K., den Ouden, E., Langerak, F., Podoynitsyna, K. (2020). Front end transfers of digital innovations in a hybrid agile-stage-gate setting, *Journal of Product Innovation Management*, Vol. 37, No. 6, 506-527, doi: 10.1111/jpim.12556.
[34] Gemino, A., Reich, B.H., Serrador, P.M. (2020). Agile, traditional, and hybrid approaches to project success: Is hybrid a poor second choice?, *Project Management Journal*, Vol. 52, No. 2, 161-175, doi: 10.1177/8756972820973082.
[35] Dakovic, M., Lalic, B., Delic, M., Tasic, N., Ciric, D. (2020). Systematic mitigation of model sensitivity in the initiation phase of energy projects, *Advances in Production Engineering & Management*, Vol. 15, No. 2, 217-232, doi: 10.14743/apem2020.2.360.

APEM journal

Advances in Production Engineering & Management
Volume 18 | Number 3 | September 2023 | pp 371–380
https://doi.org/10.14743/apem2023.3.479

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

Optimization of machining performance in deep hole boring: A study on cutting tool vibration and dynamic vibration absorber design

Li, L.[a,*], Yang, D.L.[a], Cui, Y.M.[a]

[a]School of Mechatronic Engineering, Jiangsu Normal University, Xuzhou, P.R. China

ABSTRACT

In the realm of precision engineering, particularly in deep hole boring processes, tool vibration emerges as a critical determinant of machining performance. This investigation elucidates the genesis of self-excited vibrations within deep hole boring operations and delineates the underlying mechanisms of cutting tool vibration. A focal point of this study is the optimal alignment of the boring bar to mitigate vibrational impacts, thereby enhancing surface finish quality and extending tool longevity. Central to this analysis is the employment of a Dynamic Vibration Absorber (DVA) aimed at attenuating cutting tool vibration. The deployment of DVA necessitates precise identification of modal parameters, namely the equivalent stiffness (K) and mass (M) of the cutting tool. This research juxtaposes various scholarly methodologies to amalgamate theoretical calculations with simulation approaches, thereby acquiring accurate modal parameters. Utilizing Matlab software, the vibration amplitude of the boring bar under varying spring stiffness scenarios was examined. Results indicate a direct correlation between increased stiffness and reduced amplitude, particularly when the frequency ratio g ranges between 0.5 and 1.1. Consequently, a stiffer DVA configuration is posited as more effective in vibration reduction. Furthermore, the study conducted frequency sweep experiments on a damping boring bar, utilizing a vibration excitation platform. These experiments revealed the existence of an optimal stiffness value for the DVA, thereby underscoring the significance of stiffness matching in vibration mitigation strategies.

ARTICLE INFO

Keywords:
Deep hole boring;
Boring bar;
Machining performance;
Vibration;
Dynamic vibration absorber;
Stiffness matching;
Matlab

**Corresponding author:*
6020190159@jsnu.edu.cn
(Li, L.)

Article history:
Received 20 May 2023
Revised 12 November 2023
Accepted 17 November 2023

1. Introduction

In the domain of machining operations, tool vibration has been identified as a pivotal factor impacting surface quality, material removal rates, and tool wear longevity [1-7]. Predominantly, self-excited vibrations manifest as the primary vibration type within machining contexts [8]. Chatter, a frequent occurrence, is typically initiated when the external excitation frequency aligns closely with the natural frequency of cutting tools. Intriguingly, a discrepancy between the excitation and natural frequencies, as depicted in Fig. 1, does not preclude chatter; it may arise owing to the phase differential between surface positions $x(t)$ and $x(t-T)$ at distinct temporal intervals. Variability in cutting thickness, h, during machining induces tool movement instability. Consequently, the excitation force encompasses a spectrum of frequencies, modulating in response to alterations in cutting parameters and workpiece materials [9, 10].

The underlying cause of vibration issues is often traced to inadequate damping within the structural framework. The most straightforward remedial approach involves augmenting external damping to the mechanical structure. Yet, this strategy encounters limitations due to structural and spatial constraints, rendering its application scope somewhat limited. Over the past century, the development of DVA has emerged as a viable solution for vibration reduction. Its ease of implementation and simplistic design have garnered acclaim [11]. Several scholars [12, 13] have undertaken simulation analyses to explore the impact of controlled vibration on surface roughness.

Fig. 2 illustrates the DVA, comprising a mass m, spring k, and damping c, a concept first introduced by Ormondroyd and Den Hartog in 1928 [14]. The Vibration Controlled System (VCS), consisting of mass M and spring K, integrates the DVA as an auxiliary kinetic system, thereby facilitating vibration energy absorption.

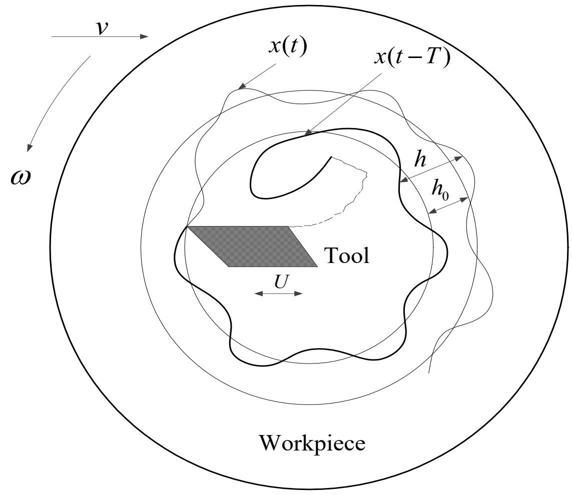


Fig. 1 Mechanism of regeneration

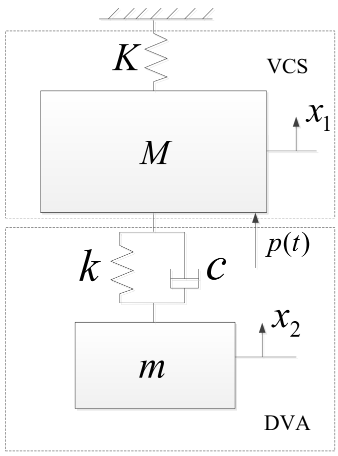


Fig. 2 Dynamic model of DVA

Prior to the design of DVA, it is imperative to accurately ascertain the equivalent stiffness (K) and mass (M) of the cutting tool. Subsequently, the parameters of DVA, i.e., mass (m), spring constant (k), and damping coefficient (c), are determined through the application of dynamic mathematical methods.

The kinetic equation of the dynamic model, as illustrated in Fig. 2, is presented as follows [15]:

$$\begin{cases} M\ddot{x}_1 + c\dot{x}_1 + (k+K)x_1 - c\dot{x}_2 - kx_2 = p(t) \\ m\ddot{x}_2 + c\dot{x}_2 + kx_2 - c\dot{x}_1 - kx_2 = 0 \end{cases} \tag{1}$$

where $p(t) = p_0 \sin \omega t$ is used to simulate external excitation signals.

Through a complex derivation process [16-19], the vibration amplitude of the VCS is derived:

$$A(g) = \sqrt{\frac{(2\zeta g)^2 + (g^2 - f^2)}{(2\zeta g)^2(g^2 - 1 + \mu g^2)^2 + [\mu f^2 g^2 - (g^2 - 1)(g^2 - f^2)]^2}} \tag{2}$$

where ζ is the damping ratio given by $\zeta = c/(2\sqrt{mk})$, μ is the mass ratio given by $\mu = m/M$, g is the frequency ratio given by $g = \omega/\omega_n$, here ω_n is the natural frequency of the main mass M given by $\omega_n = \sqrt{K/M}$, f is the natural frequency ratio given by $f = \omega_a/\omega_n$, here ω_a is the natural frequency of the DVA given by $\omega_a = \sqrt{k/m}$.

The objective is to minimize the vibration amplitude of the VCS. Upon determining the VCS parameters, the next step involves selecting optimal parameters for the DVA to achieve exemplary vibration reduction. Generally, to minimize the DVA's volume, the mass m should be set as large as possible. Given that the cutting tool body is typically composed of metal materials, the damping coefficient c is relatively small, predominantly affecting the peak height of the VCS's vibration amplitude, while exerting minimal influence on the resonance frequency. Consequently, the primary parameter of concern is the spring constant k.

Achieving optimal vibration reduction hinges on two critical factors: accurately determining the modal parameters of the cutting tool and identifying the optimal stiffness k. The modeling of the boring bar has been a subject of extensive research by several eminent scholars [20-23]. Tewani *et al*. [24] conceptualize the boring bar as an equivalent constant-section beam. Seto [25] has proposed equivalent methods for calculating the boring bar's parameters.

This study undertakes the theoretical modeling and analysis of the boring bar using three distinct methods. Subsequently, the impact of the absorber's stiffness on the amplitude of the boring bar is examined using Matlab. Finally, a frequency sweep experiment of the damping boring bar is conducted on a vibration excitation platform, yielding some noteworthy results.

2. Theoretical analysis

As depicted in Fig. 3, the boring bar functions akin to a cantilever beam during the cutting process. During this operation, the cutting head is subjected to a wide-band signal, potentially inducing multiple vibration modes within the boring bar.

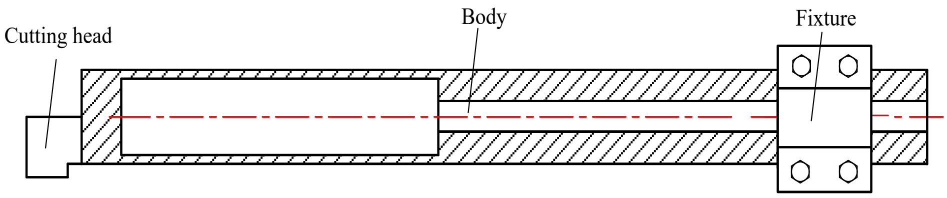


Fig. 3 Boring bar

Utilizing Abaqus software facilitates the determination of the boring bar's natural frequencies. Fig. 4 illustrates these frequencies as 350.58 Hz, 350.98 Hz, 1837.5 Hz, and 1840.1 Hz, respectively. However, the energy of high-frequency signals is comparatively weak, rarely manifesting in typical operations. Consequently, the primary focus is typically on the first-order frequency. It is evident that the lower natural frequencies are associated with bending vibrations of the boring bar.

Given the significance of the first-order bending vibration in the boring bar, the system is simplified for computational efficiency by modeling it as a one-degree-of-freedom system. This model, as illustrated in Fig. 5, is comprised of a spring and a mass, effectively capturing the essential dynamics of the boring bar's behavior.

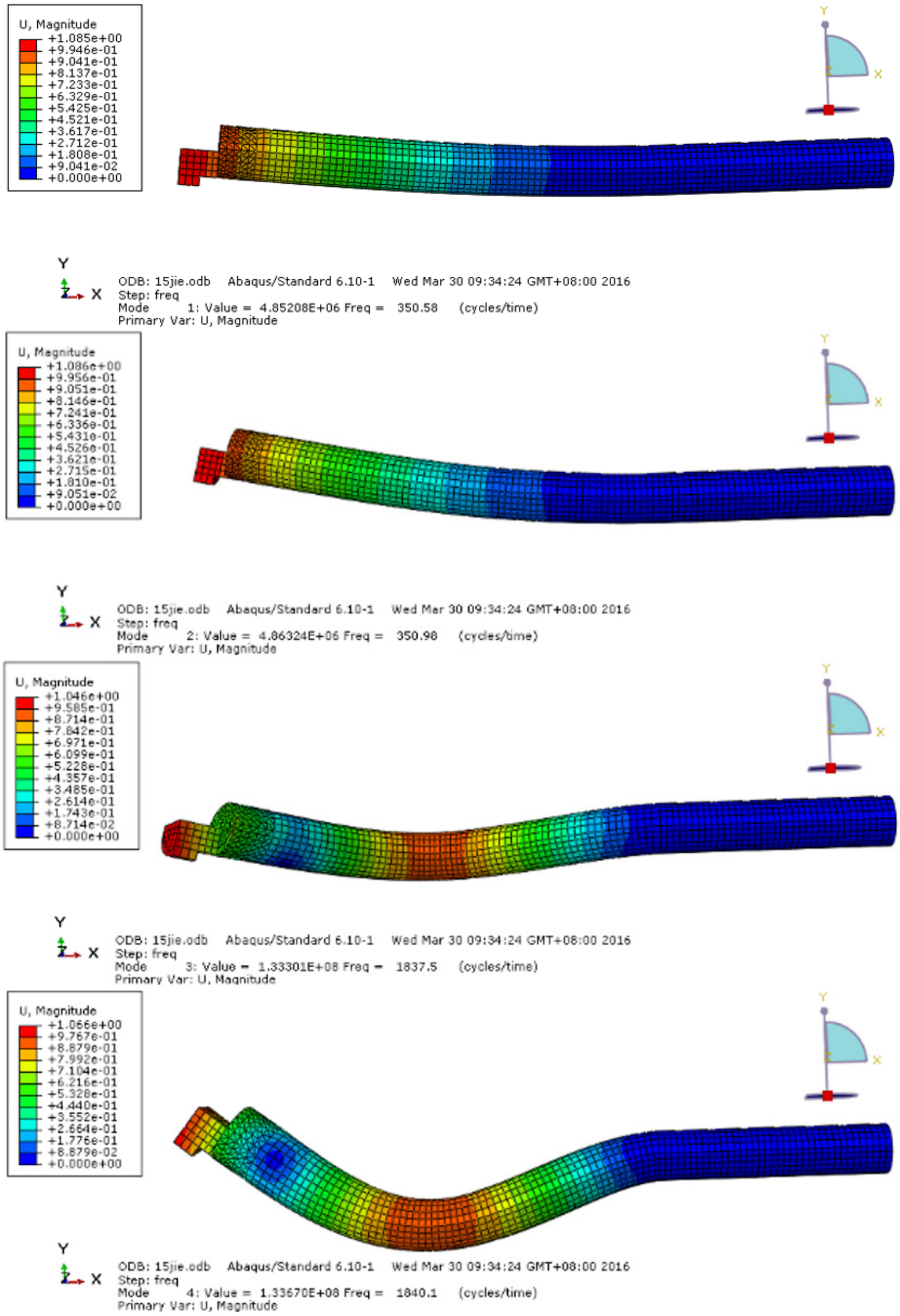


Fig. 4 Natural frequency of the boring bar

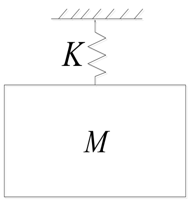


Fig. 5 Modeling of the boring bar

The equations for determining the equivalent stiffness and mass of the boring bar are as follows:

$$K = \frac{3EI}{L^3} \tag{3}$$

where L represents the length of the boring bar, E is the modulus of elasticity, and I denotes the moment of inertia.

The first-order natural frequency of a clamped beam is calculated using the equation:

$$\omega_n = 3.52\sqrt{\frac{EI}{\rho L^4}} \tag{4}$$

where ρ is the mass per unit length of the boring bar.

Then, we get the equivalent mass:

$$M = \frac{K}{{\omega_n}^2} \tag{5}$$

Apart from the aforementioned computational methods, the equivalent mass can also be estimated using an empirical formula:

$$M = m_o + 0.243m_b \tag{6}$$

where, m_o is the head mass of the boring bar, m_b is the body mass of the hollow boring bar.

Both these calculation approaches treat the boring bar as a beam with a constant cross-section. Nonetheless, for the purposes of integrating the DVA, the boring bar often features a variable cross-section. To acquire more precise values for equivalent stiffness and mass, a method that accounts for cavities in the boring bar is proposed. As illustrated in Fig. 6, a force applied at the DVA's geometric center causes a minor displacement, d_E, which can be quantified using the piece-wise superposition theory. Consequently, the equivalent stiffness is deduced as per Eq. 7:

$$K = \frac{F}{d_E} = \frac{F}{d_{E0} + d_C + \frac{1}{2}\theta_C l_2} = \frac{1}{\frac{4{l_1}^3 + 6{l_1}^2 l_2 + 3l_1{l_2}^2}{12EI_1} + \frac{F{l_2}^3}{24EI_2}} \tag{7}$$

where I_1 and I_2 are the inertia moments of AC and CD separately.

This is illustrated in Fig. 7, where the gravity acting on the boring bar induces a displacement, s, at point E. This displacement is analogous to the effect of applying a concentrated mass, M.

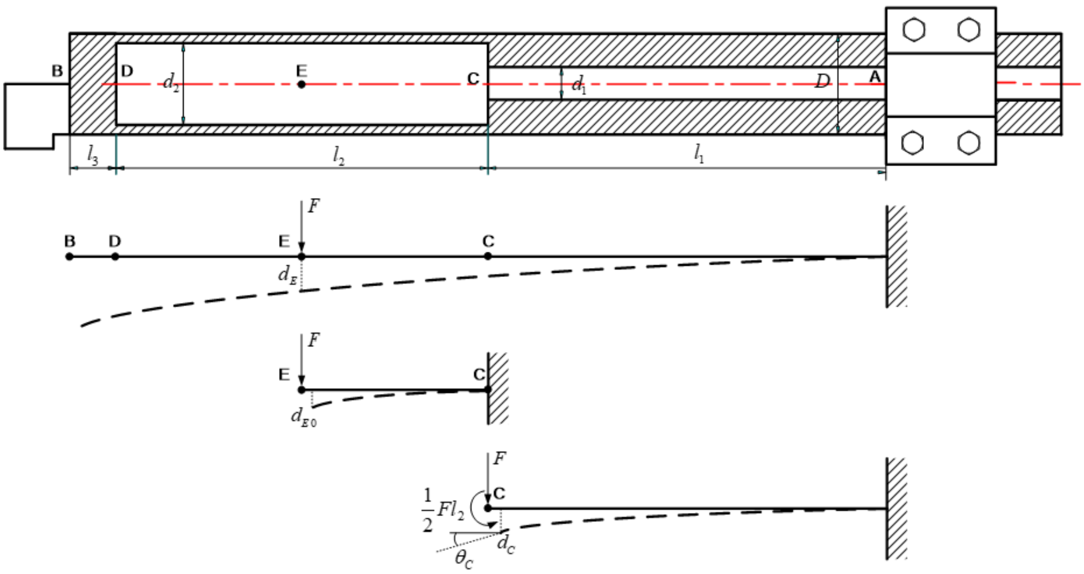


Fig. 6 Modeling of equivalent stiffness

Consequently, the equivalent mass is determined as follows:

$$M = \frac{Ks}{g} \tag{8}$$

where s is obtained:

$$\begin{aligned} s = s_1 + s_2 + s_3 = & \frac{q_1 l_1^4}{8EI_1} + \frac{q_1 l_1^3 l_2}{12EI_1} + \frac{q_2 l_1^2 l_2^2}{3EI_1} + \frac{q_2 l_1^2 l_2^2}{2EI_1} \\ & + \frac{q_2 l_1 l_2^3}{4EI_1} + \frac{17 q_2 l_2^4}{384EI_2} + \frac{q_3 l_1^3 l_3}{3EI_1} + \frac{3 q_3 l_1^2 l_2 l_3}{4EI_1} + \frac{q_3 l_1^2 l_3^2}{4EI_1} + \frac{q_3 l_1 l_2^2 l_3}{2EI_1} \\ & + \frac{q_3 l_1 l_2 l_3^2}{4EI_1} + \frac{5 q_3 l_2^3 l_3}{48EI_2} + \frac{q_3 l_2^2 l_3^2}{16EI_2} \end{aligned} \tag{9}$$

where s_1, s_2, s_3 denote the displacements of the boring bar at point E, attributable to the gravitational forces acting on different segments, as depicted in Fig. 8.

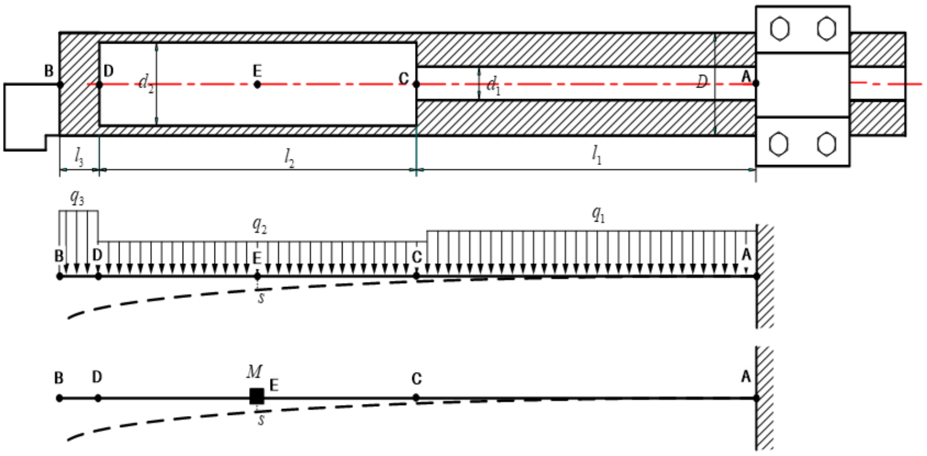


Fig. 7 Modeling of equivalent mass

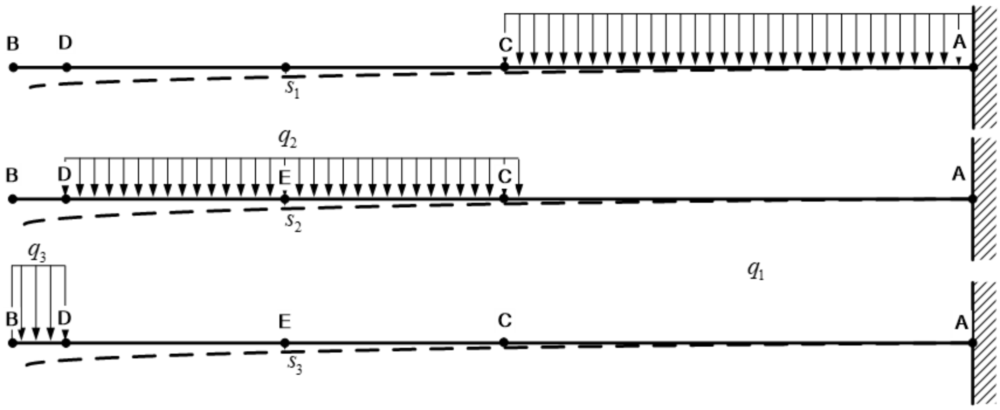


Fig. 8 Displacements under different segment's gravity

Utilizing the aforementioned three methods of calculation, results for the equivalent stiffness and mass of the boring bar have been obtained, as summarized in Table 1. In comparison with the simulation result, which indicated a natural frequency of 351 Hz, the error values have also been calculated.

Table 1 Results by using different calculation methods

Method	Equivalent stiffness (N/mm)	Equivalent mass (kg)	Frequency (Hz)	Error value (Hz)
1	2428	2.301	163.5	-186.5
2	2428	0.671	302.7	-48.3
3	2492	0.447	375.8	+24.8

It is evident that a degree of error is inherent in each calculation method used. A notable observation is the variation in equivalent mass values derived from different methods, which significantly contributes to the overall error margin. To mitigate this issue, a combined approach integrating theoretical calculations with simulation methods is employed to identify the vibration modal parameters of the boring bar more accurately. Initially, Eq. 7 is utilized to ascertain the equivalent stiffness. Subsequently, the natural frequency is determined using Abaqus, as illustrated in Fig. 4. Finally, the equivalent mass is calculated employing the equation:

$$M = \frac{K}{4\pi^2 f^2} \tag{10}$$

Fig. 9 depicts the structure of the DVA, which comprises two springs and an inner core. This assembly is filled with silicone oil. Notably, the stiffness of the DVA can be adjusted by modifying the end cap, allowing for fine-tuning of the absorber's properties to suit specific vibration control requirements.

In this study, the specific parameters under consideration are: $K = 2492$ N/mm, $M = 0.512$ kg, $m = 0.9162$ kg, $c = 237$ kg/s. Utilizing Matlab, as illustrated in Fig. 10, the vibration amplitude of the VCS is calculated across a range of spring stiffness values.

An analysis of the resultant curve reveals two distinct peaks. It is observed that as the stiffness increases, the amplitude of the first peak also rises, which is contrary to the desired outcome. However, it is important to note that the machine's excitation frequency component typically exceeds 200 Hz. This implies that the damping boring bar becomes effective when the parameter g is greater than 0.57. As depicted in Fig. 10, within the range of 0.5 to 1.1 for the parameter g, an increase in stiffness correlates with a reduction in amplitude. Therefore, to achieve more effective vibration reduction, a preference for higher stiffness is suggested.

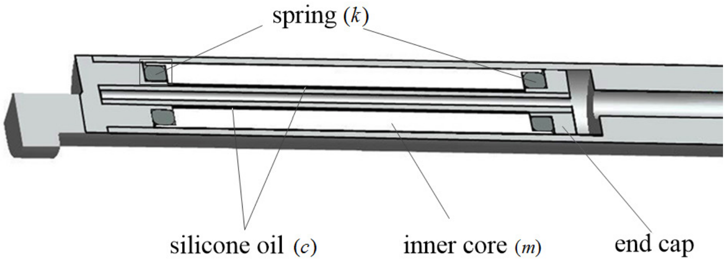


Fig. 9 Structure of the DVA

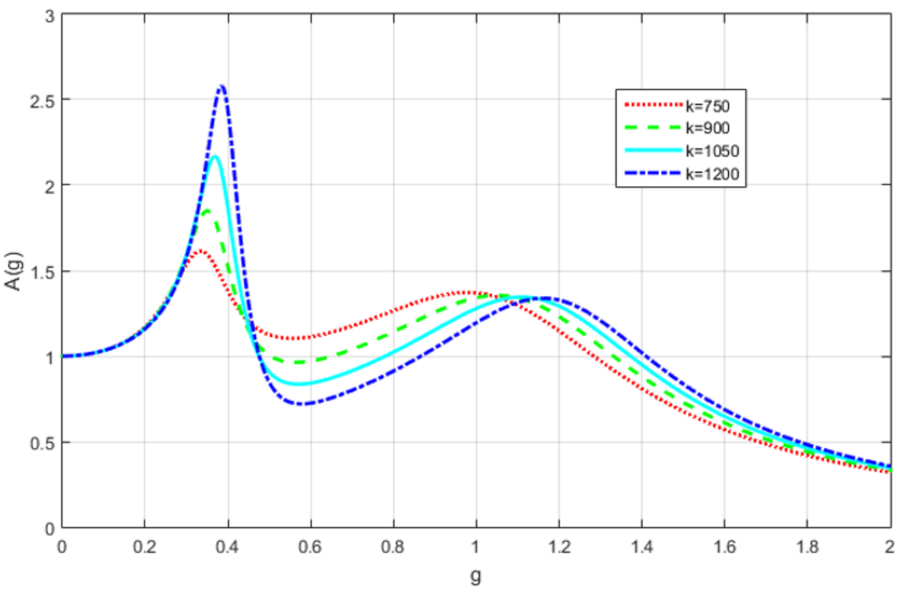


Fig. 10 Vibration amplitude of the VCS under different spring stiffness

3. Experimental work

To corroborate the analytical results, a frequency sweep experiment was conducted on a vibration excitation platform, as depicted in Fig. 11. This platform can generate a sine signal with a frequency range spanning from 20 Hz to 1200 Hz. The input control channel, responsible for maintaining the amplitude of the input signal at 1g, is strategically positioned at the tail end of the boring bar.

The experiment was conducted under four distinct conditions, paralleling the simulations previously described.

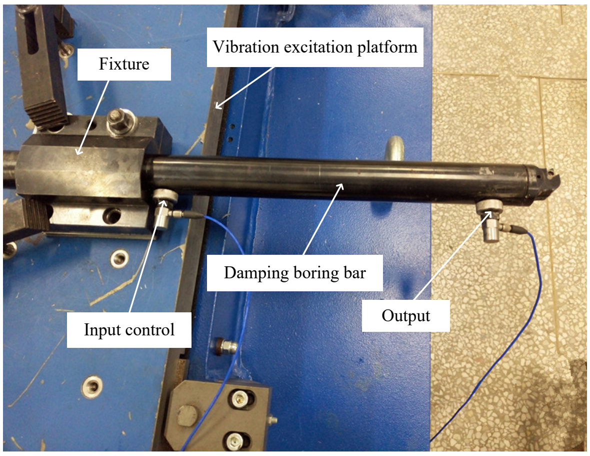


Fig. 11 Frequency sweep experiment

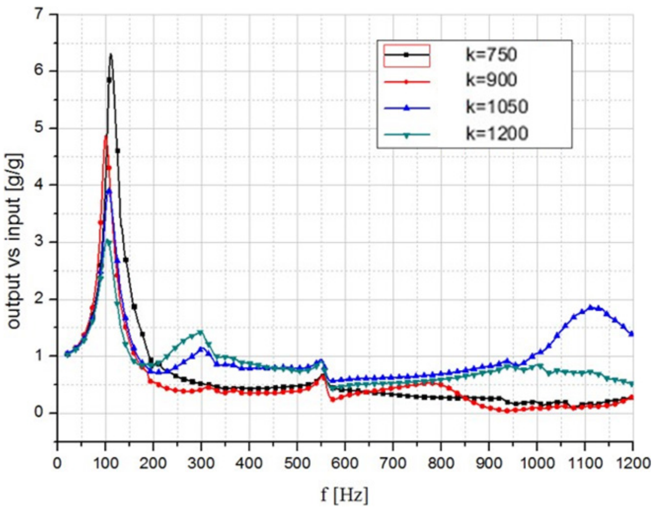


Fig. 12 Harmonic response under four different conditions

Fig. 12 presents the results of the experiment, and when compared with Fig. 10, the trends in both figures are observed to be consistent. The following key observations can be made: (1) While the spring stiffness does affect the first-order natural frequency, this impact is relatively subtle, with recorded frequencies at 110 Hz, 100 Hz, 105 Hz, and 103 Hz, respectively; (2) An increase in stiffness leads to a higher peak amplitude of the first-order natural frequency; (3) As previously emphasized, the primary focus is on the amplitude when the excitation frequency exceeds 200 Hz. In the context of the four experimental conditions detailed in this study, the lowest vibration amplitude of the boring bar is achieved with a spring stiffness of 900 N/m. However, this finding deviates from the earlier simulation results. This discrepancy can be attributed to the fact that at very high stiffness levels, the inner core and the boring bar behave as a single entity, thereby rendering the damping boring bar no longer a two-degree-of-freedom system but a more complex dynamic system.

4. Conclusion

Vibration is an inherent phenomenon in machining processes. The DVA is recognized as an effective solution for vibration mitigation, owing to its simplicity in implementation and structure. To optimize the vibration reduction efficacy of the DVA, two critical aspects must be addressed: accurately determining the modal parameters of the cutting tool, and identifying the optimal spring stiffness, k. A combination of theoretical calculations and simulation methods is employed to ascertain these modal parameters, including the equivalent stiffness (K) and mass (M). Simulations conducted using Matlab reveal the relationship between spring stiffness and the vibration amplitude of the VCS. The results indicate an increase in the peak amplitude of the first-order natural frequency with higher stiffness. However, within the parameter g range of 0.5 to 1.1, a decrease in amplitude is observed, suggesting that a larger stiffness is preferable for improved vibration reduction.

Experimental validation was performed through frequency sweep experiments on a damping boring bar using a vibration excitation platform. It was found that exceedingly high stiffness levels cause the inner core and the boring bar to function as a single unit, thereby altering the system's dynamics from a two-degree-of-freedom to a more complex state. Consequently, excessively high stiffness does not yield optimal vibration reduction. Thus, it is imperative to select a stiffness level that is appropriate for the specific machining model to achieve the best vibration reduction outcome.

Acknowledgement

This work is supported by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 20KJB460018), the Research Fund for Doctoral Degree Teachers of Jiangsu Normal University, China (Grant No. 20XSRS014), the National Natural Science Foundation of China (Grant No. 12302013 Study on nonlinear dynamic characteristics and vibration suppression mechanism of an air-supported boring bar), the Practice Innovation Training Program for College Students, China (Grant No. 202110320087Y).

References

[1] Quintana, G., Ciurana, J. (2011). Chatter in machining processes: A review, *International Journal of Machine Tools and Manufacture*, Vol. 51, No. 5, 363-376, doi: 10.1016/j.ijmachtools.2011.01.001.

[2] Okokpujie, I.P., Sinebe, J.E. (2023). An overview of the study of ANN-GA, ANN-PSO, ANFIS-GA, ANFIS-PSO and ANFIS-FCM predictions analysis on tool wear during machining process, *Journal Européen des Systèmes Automatisés,* Vol. 56, No. 2, 269-280, doi: 10.18280/jesa.560212.

[3] Siddhpura, M., Paurobally, R. (2012). A review of chatter vibration research in turning, *International Journal of Machine Tools and Manufacture,* Vol. 61, 27-47, doi: 10.1016/j.ijmachtools.2012.05.007.

[4] Antonucci, A., Coltrinari, G., Lippiello, D. (2023). Effectiveness of antivibration gloves when used with a light electric hammer, Differences among different methods of measurements, *International Journal of Computational Methods and Experimental Measurements*, Vol. 11, No. 1, 27-34, doi: 10.18280/ijcmem.110104.

[5] Margabandu, V., Radhakrishnan, R. (2021). Multi objective study on machining characteristics of AISI H-11 tool steel prepared by different processing techniques, *Journal Européen des Systèmes Automatisés,* Vol. 54, No. 2, 243-251, doi: 10.18280/jesa.540206.

[6] Okokpujie, I.P., Tartibu, L.K. (2022). Comparative study of the effect of dry, mineral oil, and TiO2 nano-lubricant on tool wear during face-milling machining of Ti-6al-4v-Eli using carbide tool insert, *Mathematical Modelling of Engineering Problems,* Vol. 9, No. 2, 468-476, doi: 10.18280/mmep.090224.

[7] Gao, H.N., Shen, H.D., Yu, L., Wang, Y.L., Yang, Y., Yan, S.C., Hu, Y.J. (2021). Frictional wear detection of hard alloy tool material during high-speed cutting, *International Journal of Heat and Technology,* Vol. 39, No. 6, 1845-1852, doi: 10.18280/ijht.390619.

[8] Jasiewicz, M., Miądlicki, K. (2020). An integrated CNC system for chatter suppression in turning, *Advances in Production Engineering & Management*, Vol. 15, No. 3, 318-330, doi: 10.14743/apem2020.3.368.

[9] Li, L., Sun, B., He, M. (2019). Analysis of radial stiffness of rubber bush used in dynamic vibration absorber, *Journal of Southeast University (English Edition)*, Vol. 35, No. 3, 281-287.

[10] Vukelic, D., Prica, M., Ivanov, V., Jovicic, G., Budak, I., Luzanin, O. (2022). Optimization of surface roughness based on turning parameters and insert geometry, *International Journal of Simulation Modelling*, Vol. 21, No. 3, 417-428, doi: 10.2507/IJSIMM21-3-607.

[11] Rubio, L., Loya, J.A., Miguélez, M.H., Fernández-Sáez, J. (2013). Optimization of passive vibration absorbers to reduce chatter in boring, *Mechanical Systems and Signal Processing*, Vol. 41, No. 1-2, 691-704, doi: 10.1016/j.ymssp.2013.07.019.

[12] Kang, W.T., Derani, M.N., Ratnam, M.M. (2020). Effect of vibration on surface roughness in finish turning: Simulation study, *International Journal of Simulation Modelling*, Vol. 19, No. 4, 595-606, doi: 10.2507/IJSIMM19-4-531.

[13] Vukelic, D., Simunovic, K., Kanovic, Z., Saric, T., Doroslovacki, K., Prica, M., Simunovic, G. (2022). Modelling surface roughness in finish turning as a function of cutting tool geometry using the response surface method, Gaussian process regression and decision tree regression, *Advances in Production Engineering & Management*, Vol. 17, No. 3, 367-380, doi: 10.14743/apem2022.3.442.

[14] Ormondroyd, J., Den Hartog, J.P. (1928). The theory of the dynamic vibration absorber, *ASME Journal of Applied Mechanics,* Vol. 50, No. 7, 9-22.

[15] Miguélez, M.H., Rubio, L., Loya, J.A., Fernández-Sáez, J. (2010). Improvement of chatter stability in boring operations with passive vibration absorbers, *International Journal of Mechanical Sciences*, Vol. 52, No. 10, 1376-1384, doi: 10.1016/j.ijmecsci.2010.07.003.

[16] Li, L., Sun, B. (2016). Optimal parameters selection and engineering implementation of dynamic vibration absorber attached to boring bar, In: *Proceedings of the INTER-NOISE and NOISE-CON Congress and Conference,* Hamburg, Germany, 563-570.

[17] Li, L., Sun, B., Hua, H. (2019). Nonlinear system modeling and damping implementation of a boring bar, *International Journal of Advanced Manufacturing Technology*, Vol. 104, 921-930, doi: 10.1007/s00170-019-03907-8.

[18] Li, L., Sun, B., Hua, H. (2019). Analysis of the vibration characteristics of a boring bar with a variable stiffness dynamic vibration absorber, *Shock and Vibration*, Vol. 2019, Article ID 5284194, doi: 10.1155/2019/5284194.

[19] Brock, J.E. (1946). A note on the damped vibration absorber, *ASME Journal of Applied Mechanics,* Vol. 13, No. 4, Article No. A-284, doi: 10.1115/1.4009588.

[20] Åkesson, H., Smirnova, T., Håkansson, L. (2009). Analysis of dynamic properties of boring bars concerning different clamping conditions, *Mechanical Systems and Signal Processing*, Vol. 23, No. 8, 2629-2647, doi: 10.1016/j.ymssp.2009.05.012.

[21] Andrén, L., Håkansson, L., Brandt, A., Claesson, I. (2004). Identification of dynamic properties of boring bar vibrations in a continuous boring operation, *Mechanical Systems and Signal Processing*, Vol. 18, No. 4, 869-901, doi: 10.1016/S0888-3270(03)00093-1.

[22] Zhang, G.M., Kapoor, S.G. (1987). Dynamic modeling and analysis of the boring machining system, *ASME Journal of Engineering for Industry,* Vol. 109, No. 3, 219-226, doi: 10.1115/1.3187122.

[23] Parker, E.W. (1970). Dynamic stability of a cantilever boring bar with machined flats under regenerative cutting conditions, *Journal of Mechanical Engineering Science*, Vol. 12, No. 2, 104-115, doi: 10.1243/JMES_JOUR_1970_012_018_02.

[24] Tewani, S.G., Rouch, K.E., Walcott, B.L. (1995). A study of cutting process stability of a boring bar with active dynamic absorber, *International Journal of Machine Tools and Manufacture*, Vol. 35, No. 1, 91-108, doi: 10.1016/0890-6955(95)80009-3.

[25] Seto, K., Yamada, K. (1980). An investigation on boring bars equipped with a dynamic absorber, In: *Proceeding of the 4th International Conference on Production Engineering*, Tokyo, Japan, 422-427.

APEM journal

Advances in Production Engineering & Management
Volume 18 | Number 3 | September 2023 | pp 381–395
https://doi.org/10.14743/apem2023.3.480

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

Optimal logistics scheduling with dynamic information in emergency response: Case studies for humanitarian objectives

Cao, J.[a,*], Han, H.[b], Wang, Y.J.[c], Han, T.C.[d]

[a]Xuzhou University of Technology, Xuzhou, Jiangsu, P.R. China
[b]Shanxi College of Technology, School of Information Industry, Shuozhou, Shanxi, P.R. China
[c]Shuozhou Meteorological Bureau of Shanxi Province, Shuozhou, Shanxi, P.R. China
[d]London School of Economics and Political Science, London, United Kingdom

A B S T R A C T

The mathematical model of infectious disease is a typical problem in mathematical modeling, and the common infectious disease models include the susceptible-infected (SI) model, the susceptible-infected-recovered model (SIR), the susceptible-infected-recovered-susceptible model (SIRS) and the susceptible-exposed-infected-recovered (SEIR) model. These models can be used to predict the impact of regional return to work after the epidemic. In this paper, we use the SEIR model to solve the dynamic medicine demand information in humanitarian relief phase. A multistage mixed integer programming model for the humanitarian logistics and transport resource is proposed. The objective functions of the model include delay cost and minimum running time in the time-space network. The model describes that how to distribute and deliver medicine resources from supply locations to demand locations with an efficient and lower-cost way through a transportation network. The linear programming problem is solved by the proposed Benders decomposition algorithm. Finally, we use two cases to calculate model and algorithm. The results of the case prove the validity of the model and algorithm.

A R T I C L E I N F O

Keywords:
Logistic;
Humanitarian logistics;
Optimization;
Multi-objective;
Dynamic information;
Delay cost;
Benders decomposition algorithm;
Mixed integer programming;
Ant colony optimization algorithm;
Genetic algorithm

**Corresponding author:*
cj@nuist.edu.cn
(Cao, J.)

Article history:
Received 15 December 2022
Revised 9 September 2023
Accepted 15 September 2023

1. Introduction

In recent years, a variety of natural disasters occur frequently in the world, e.g., a 7.1 earthquake injured 12 thousand people in Qinghai province in 2010, a 7.0 earthquake occurred in Ya'an city in April 2013, etc. They have catastrophic effects on the society and everyone's daily life in many aspects such as injuries, property damage and even loss of life [1]. How to effectively respond to unpredictable and irregular emergency events such as dynamic resource allocation has become of primal importance worldwide [2].

Epidemic events usually follow major natural disasters and cause secondary damage to the people in the disaster areas. For example, dysentery, measles, pinkeye and epidemic encephalitis B may break out after an earthquake. Once an epidemic breaks out, it will cause incalculable losses. Therefore, it remains imperative for every country to be prepared for emergency rescue if an infectious disease breaks out. Ensuring the supply of medical resources would always be the most important factor in the prevention and treatment of the epidemic diffusion.

Emergency medicine resource allocation is aimed to aid people and allocate relief distribution with dynamic information in surviving during and after an infectious disease occurs. The public officials need to face and solve many critical issues, the most important of which is how to timely and reasonably distribute the emergency medicine resources to the epidemic area so that the delay cost [3] of rescue medicine resources can be minimised and the rescue vehicle running time can be minimised. The delay cost of rescue medicine resources can be regarded as the standard to capture the social costs of shortage of rescue resources, which is called deprivation costs, has been roundly proposed and classified in Holguin-Veras *et al.* [4]. Due to the complexity and uncertainty of the emergency logistics, material supply and distribution is more difficult. First of all, it is difficult to obtain the demand information of emergency material resources, especially for the epidemics with uncertain infectivity. Second, rescue time is a critical factor after medical emergency events occur, the delay of material supply can cause inestimable loss. Finally, managers and decision makers need to distribute the emergency materials to affected areas as quickly and effectively as possible.

In this paper, a time-space network model for the humanitarian logistics problem with logistics in controlling epidemic diffusion is proposed. It couples a forecasting mechanism for the number of people who need to treat based on the SEIR model [5] and a multistage programming model for the humanitarian logistics and transport resource. Our contribution includes that the dynamic demand of the medical resource based on the epidemic diffusion pattern of SEIR model is used in a multistage programming model for optimal allocation and transport of such resource and the linear programming problem is solved by the Benders decomposition algorithm which is better than ant colony algorithm and genetic algorithm in computational accuracy.

Our paper is organized as follows. In Section 2, we review the literature relevant to our study. In Section 3, we build the time-space network model, which combines a demand forecast model based on the epidemic diffusion rule. The solution procedure for the optimization model is proposed in Section 4. In Section 5, we report the results of computational studies and sensitive analysis. Finally, we discuss the conclusions and suggest future research directions in Section 6.

2. Literature review

Our contribution is connected with four research branches: 1) the mathematical formulations of the epidemic diffusion mechanism; 2) the multistage programming model on the humanitarian logistics and transport resource; 3) the combination of medical model and humanitarian logistics model: 4) the solution methodology, such as Benders decomposition algorithm, genetic algorithm and ant colony algorithm.

There are many analytical works on epidemic diffusion including SIR epidemic models, Susceptible-Infected-Susceptible (SIS) epidemic model, SIRS epidemic model, SEIR epidemic model, susceptible-exposed-infected-recovered-susceptible (SEIRS) epidemic model and so on. In these models, most of them are developed by ordinary differential equations. To develop the epidemic diffusion mechanism, Li *et al.* [6] study the spread dynamics of a stochastic SIRS epidemic model with nonlinear incidence and varying population size, which is formulated as a piece wise deterministic Markov process. Fan *et al.* [7] present a SIR epidemic model with generalized nonlinear incidence rate. Song *et al.* [8] propose a SEIR reaction-diffusion model, where the disease transmission and recovery rates can be spatially heterogeneous. Yang and Wang [9] introduce a new SEIRS epidemic model with time delay on a scale-free network.

Some of these models are used in the complex and actual problems. Bolzoni *et al.* [10] investigate the time-optimal control problem in SIR epidemic models, focusing on different control policies: vaccination, isolation, culling, and reduction of transmission. Guo *et al.* [11] explore the global behaviour of a stochastic SIRS epidemic model with media coverage. Britton and Ouédraogo [12] introduce an SEIRS epidemic with disease fatalities in a growing population. Anparasan and Lejeune [13] propose an epidemic response model in resource-limited countries that determines the number, size, and location of treatment facilities, deploys critical medical staff, locates ambulances to triage points, and organizes the transportation of severely ill patients to treatment facilities. The SEIR model is used to solve the dynamic medicine demand information in humanitarian relief phase. These models can be used to predict the impact of regional return to work after the epidemic. There are other forecasting methods, which are widely used in other fields, such as the artificial neural network [14], the artificial intelligence [15], and so on.

To deal with the complexity and difficulty in solving the humanitarian logistics and transport resource scheduling problem, we summarize the research for the humanitarian logistics and transport resource scheduling problem in recent years. The literature review of humanitarian logistics summarizes the research status in this field. Farahani *et al.* [16] summarize the mass casualty management including five steps: (i) Resource dispatching/search and rescue, (ii) on-site triage, (iii) on-site medical assistance, (iv) transportation to hospitals and (v) triage and comprehensive treatment. Baffoe and Luo [17] use a systematic literature review coupled with an axiological philosophical lens approach to developing a Humanitarian Logistics Digital Business Ecosystem (HLDBE) framework as an alternative way to sustain the humanitarian logistics operations and reliefs through hybrid humanitarian- business logistics sector. Wang *et al.* [18] study the routing problem of unmanned vehicles considering path flexibility. Anbal *et al.* [19] use a new technology to solve the intelligent traffic management. Wang *et al.* [20] discusses the model of joint distribution of fast-moving consumer goods.

To optimize the process of humanitarian logistic problem, Zhou *et al.* [21] design a multi-objective optimization model for multi-period dynamic emergency resource scheduling (ERS) problems. Othman *et al.* [22] propose a multi-agent-based architecture for the management of Emergency Supply Chains (ESCs), in which each zone is controlled by an agent. A Decision Support System (DSS) states and solves, in a distributed way, the scheduling problem for the delivery of resources from the ESC supplying zones to the ESC crisis-affected areas. Ferrer *et al.* [23] build a compromise programming model for multi-criteria optimization in humanitarian last mile distribution and illustrate the multi-criteria optimization using a realistic test case based on the Pakistan floods, 2010.

To deal with the efficiency and timeliness in solving the humanitarian objectives, Huang *et al.* [2] characterize the humanitarian objectives of emergency resource allocation and distribution in disaster response operations. And they formulated the humanitarian principles as three objective functions, i.e., lifesaving utility, delay cost and fairness. Garrido and Aguirre [24] present a modelling framework to assist decision makers in strategic and tactical planning for effective relief operations after an earthquake's occurrence. The objective is to perform these operations quickly while keeping its total expenses under a budget. Edward *et al.* [25] consider a joint vehicle and crew routing and scheduling problem in which crews are able to interchange vehicles, resulting in space and time interdependencies between vehicle routes and crew routes. Rajak *et al.* [26] present a hybrid metaheuristic which combines simulated annealing, ant colony optimization and along with long-arc-broken removal heuristic approach for solving the multi-depot vehicle routing problem with simultaneous deliveries and pickups. Sedehzadeh and Deifbarghy [27] present a closed loop food supply chain network. The objectives of the model are to minimize costs, transportation emissions, and unsatisfied foodbanks' demand. Foodbank as the main pillar of social responsibility has been introduced in food chain.

In order to solve the dynamic demand information of emergency material resources in rescue process. Gutjahr and Nolz [28] review recent literature on the application of multicriteria optimization to the management of natural disasters, epidemics or other forms of humanitarian crises. Liu and Xiao [29] model for a dynamic resource allocation problem following an epidemic outbreak in a region. Liu [30] develop a unique time-varying forecasting model for dynamic demand of medical resources based on a SEIR influenza diffusion model. Wang *et al.* [1] construct a multi-objective stochastic programming model with time-varying demand for the emergency logistics network based on the epidemic diffusion rule. Buschiazzo *et al.* [31] consider stockouts costs and inventory maintenance costs in their model for healthcare supplies problem.

To deal with the mixed integer programming problem, Fischetti *et al.* [32] prove that Benders decomposition allows for a significant boost in the performance of a mixed-integer programming solver. And in order to improve computation efficiency, authors investigate the use of proximity search as a tactical tool to drive Benders decomposition. Alkaabneh *et al.* [33] consider the problem of inventory routing in the context of perishable products and find near-optimal replenishment scheduling and vehicle routes, and develop an exact method based on Benders decomposition to find high-quality solutions in reasonable time. Fachini and Armentano [34] present exact algorithms based on logic-based Benders decomposition and a variant, called branch- and- check, for the heterogeneous fixed fleet vehicle routing problem with time windows. Cordeau *et al.* [35] study an effective decomposition approach to the two problems based on the branch-and-Benders-cut reformulation. The proposed approach is designed for the realistic case in which the number of customers is much larger than the number of potential facility locations. Behmanesh and Rahimi [36] use the ant colony optimization to solve the multi-resource job shop scheduling problem. Fei [37] proposes intelligent bionic optimization algorithm based on the growth characteristics of tree branches. The mixed integer programming problem can be solved by other heuristic algorithms, such as genetic algorithms. Although genetic algorithm can improve the computational efficiency, it is often unable to obtain accurate results and is easy to fall into local optimal solutions. When the scale of the problem is large, genetic algorithm is a good choice.

Furthermore, we note that most of the previous epidemic models were innovated by developing differential equation and most of the resource allocation in the humanitarian logistics rarely take into the dynamic demand information. In addition, only a few literatures combine epidemic models with humanitarian logistics. While in reality, the demand for medical resource is dynamic, and the medical resource allocated in early cycles will affect the demand in later periods [29]. In this paper, a novel SEIR epidemic model is used to forecast the time-varying demand in humanitarian logistics. We use a time-space network to describe the humanitarian logistics when an epidemic occurs. In each decision cycle, the problem is constructed as a linear programming model to solve for the delay cost minimizing and vehicle running time minimizing. A Benders decomposition algorithm is compared with other heuristic algorithms in solving humanitarian logistics with dynamic demand information.

3. The mathematical model

3.1 Problem description

A disaster often causes epidemics, i.e., cholera, typhoid fever, dysentery often ravage disaster area after the floods. Therefore, it is important that government and social organization send medicine and vaccine to the disaster area in time. A large amount of distribution costs can be generated when the drug is delivered. Decision makers need follow the low-budget principle in the delivery process. Our research problem is how to distribute and deliver medicine resources from supply locations to demand locations with an efficient and lower-cost way through a transportation network.

Due to the complexity and unpredictability of disasters and the property of the epidemics, as time goes on, demand of infectious disease patients is changed. In our paper, dynamic demand is defined as the number of people requiring treatment. Firstly, we use SEIR model to calculate the number of infected people. Then the demand of infected people in disaster area can be simulated in the forecasting model for the time-varying demand. Finally, we construct time-space network of the humanitarian logistics model to optimize delay cost and transportation cost functions.

In our model formulation, we use a time-space network to describe entire delivery process. The time-space network of the humanitarian logistics is shown in Fig. 1. In the time-space network, decision cycle is divided into several discrete time units $t = 0, 1, 2, \dots, \mathrm{T}$. t represents the decision point for the decision cycle. There are three arcs in the time-space network: (1) holding arc, i.e., arc (a); (2) allocation arcs, i.e., arc (b), which represents that vehicle originates at a distribution center at time t and arrives at a demand point at time t', $0 \leq t < t' \leq$ T, t'- t is the vehicle running time on the arc; arc (c), which represents that vehicle originates at a demand point at time t and arrives at a demand point at time t', $0 \leq t < t' \leq$ T; (3) return arc, i.e., arc (d), which represents that vehicle originates at a demand point at time t and arrives at a distribution center at time t', $0 \leq t < t' \leq$ T.

Let D denote the collection of all distribution centers and N denotes the collection of all nodes. If node $i, j \in D$, node i and j represent distribution center, otherwise, node i and j represent demand point. We use ij to denote the arc in geography network. Let A denote the collection of arcs in geography network. Let k represent the vehicle number respectively. Let K denote the collection of vehicles. We use Q_k to represent the maximum loading capacity of the vehicle k. α denotes the cost conversion coefficient.

There are two types of decision variables in the whole rescue vehicle dispatch period. Let $y^k_{it,jt'}$ represent the flow originating at node i at time t and arriving at node j at time t'. $y^k_{it,jt'}$ is a binary decision variable. We view $y^k_{it,jt'}$ as the vehicle selection decision variables. There is an arc (it, it'), which allows the vehicle resource to stay at the same node from time t to time t'. We use N^T and A^T to denote the collections of nodes and arcs in the time space network. Let x^k_{jt} represent the resource allocated to node j and received at time t by vehicle k, and x^k_{jt} is the distribution decision variable.

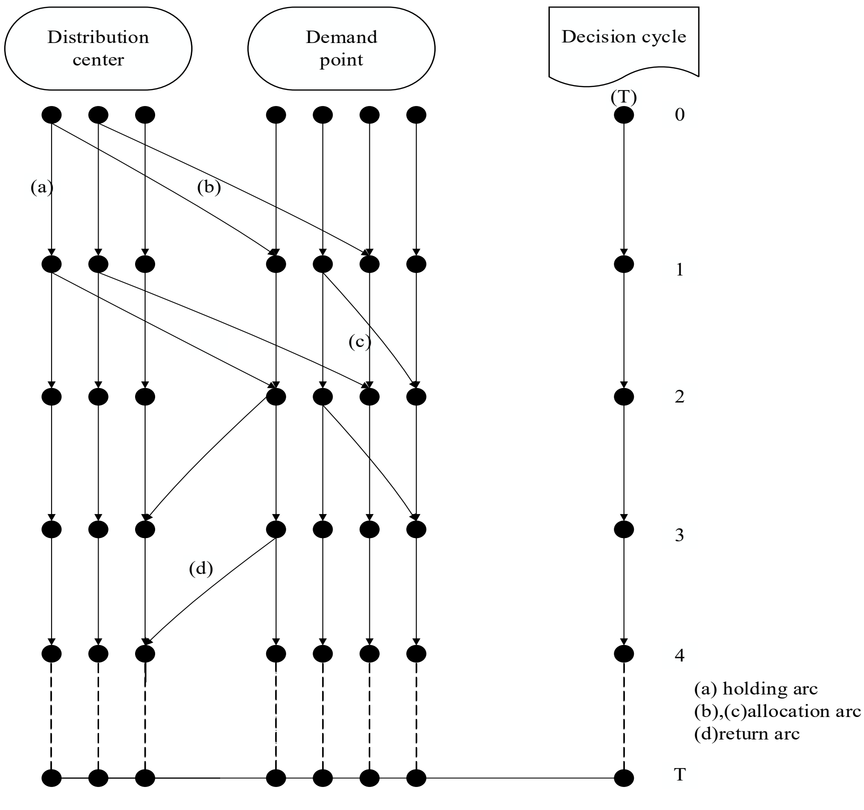


Fig. 1 Time-space network of the humanitarian logistics

3.2 The forecasting model for the time-varying demand

To represent the demand, supply and demand locations, we use the term 'node' uniformly, and use i and j as index. We use $P^t_j \geq 0$ to represent the quantity of unsatisfied demand of the epidemic patients at node j updated at decision point t. In order to calculate the demand of the epidemic patients, we use SEIR model that can simulate susceptible people (S), exposed people (E), infected people (I) and recovered people (R) in disaster area to obtain the number of infected people at node j at time t. We use S_j (t), E_j (t), I_j (t), and R_j (t) to represent, respectively, the number of susceptible people at node j at time t, the number of exposed people at node j at time t, the number of infected people at node j at time t, and the number of recovered people at node j at time t. Fig. 2 shows, without consideration of migration, the natural birth rate and death rate of the population, the epidemic process can be described by a SEIR model based on a small-world network.

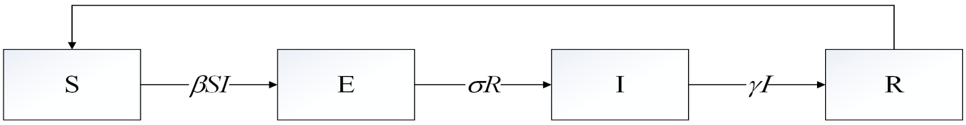


Fig. 2 SEIR model based on a small-world network.

The dynamic system for the SEIR diffusion model is modified based on the literature [38]. It can be rewritten by the following differential equations:

$$\begin{gathered} \frac{dS_j}{d_t} = -\beta S_j(t) I_j(t), \\ \frac{dE_j}{d_t} = S_j(t) I_j(t) - \sigma R_j(t), \\ \frac{dI_j}{d_t} = \sigma R_j(t) - \gamma I_j(t), \\ \frac{dR_j}{d_t} = \gamma I_j(t), \end{gathered} \tag{1}$$

where β is contact rate (Susceptible to Exposed). σ is incubation rate (Exposed to Infected). γ is recovery rate (Infected to Recovered).

In this study, $I_j(t)$ denote the number of people who need to treat. Therefore, P_j^t with time-varying can be rewritten by:

$$P_j^t = \lambda I_j(t) \tag{2}$$

where λ denotes the proportionality coefficient. λ in this linear forecasting function is adopted by the public-healthcare administrative personnel in controlling the spread of epidemic [2]. Herein, we don't think about the lag effect of earlier medicine allocation.

3.3 Time-space network of the humanitarian logistics

We consider two objectives related to humanitarian logistics: time cost and delay cost. $\sum_{it,jt' \in A^T} \sum_{k \in K} (t' - t) y_{it,jt'}^{\mathrm{k}}$ denotes the total time cost. The accumulated number of medical sources have been delivered by vehicles at node j at time t can be rewritten by $\sum_{k \in K} \sum_{t=0}^{t} x_{jt}^k$. Let $P_j^t - \sum_{k \in K} \sum_{t=0}^{t} x_{jt}^k$ denote the unsatisfied demands at node j at time t. The total delay cost can be rewritten by $\sum_{j \in N} \sum_{t \in T} (P_j^t - \sum_{k \in K} \sum_{t=0}^{t} x_{jt}^k)$. The time-space network of the humanitarian logistics can be rewritten by:

$$Minimize \ \sum_{j \in N} \sum_{t \in T} (P_j^t - \sum_{k \in K} \sum_{t=0}^{t} x_{jt}^k) + \alpha \sum_{it,jt' \in A^T} \sum_{k \in K} (t' - t) y_{it,jt'}^{\mathrm{k}} \tag{3}$$

$$s.t. \ \sum_{it \in N^T} y_{it,jt'}^k = 1, \forall jt' \in N^T, k \in K, \tag{4}$$

$$\sum_{jt' \in N^T} y_{it,jt'}^k = 1, \forall it \in N^T, k \in K, \tag{5}$$

$$\sum_{ij \in A} y_{it,jt'}^k = 1, \forall t, t' \in T, k \in K, \tag{6}$$

$$\sum_{it \in N^T} \sum_{k \in K} y_{it,jt'}^k = \sum_{it \in N^T} \sum_{k \in K} y_{jt',it}^k, \forall jt' \in N^T, \tag{7}$$

$$\sum_{\forall jt' \in N^T} x_{jt'}^k \le Q_k \times \sum_{it,jt' \in A^T} y_{it,jt'}^k, \forall k \in K, \tag{8}$$

$$\sum_{k \in K} x_{jt}^k \le P_j^t, \forall t \in T, j \in N, \tag{9}$$

$$y_{it,jt'}^k \in \{0,1\}, \forall it, jt' \in N^T, k \in K, \tag{10}$$

$$x_{jt}^k \ge 0, integer, \forall jt \in N^T, k \in K. \tag{11}$$

In this optimization model, the objective function (Eq. 3) is to minimize the total cost of humanitarian logistics. Constraints Eqs. 4-7 are the flow conservation equations. Particularly, constraint (Eq. 4) constraint (Eq. 5) and Constraint (Eq. 6) ensure that each node must be visited in the time-space network. Constraint (Eq. 7) states that each vehicle must get away of the node after it arrives one node in the time-space network. Constraint (Eq. 8) ensures that each vehicle load number of people is smaller and equal to its maximum load at node j. Constraint (Eq. 9) states that the number of delivered medicine resource is smaller and equal to the demand number at node j at time t. Finally, Constraint (Eq. 10) ensures that $y_{it,jt'}^k$ is binary variable. Constraint (Eq. 11) are the nonnegativity of the flows. Such model is a dynamic and multistage programming model.

4. Solution methodology

To solve the above optimization model, Eq. 1 and Eq. 2 are adopted to calculate the time-varying demand P_j^t firstly. After that, we use Benders decomposition algorithm to solve the mixed integer programming model. Benders decomposition algorithm provides a basic framework to solve MILP through decomposing the original complex problem into two problems, i.e., a master problem and a subproblem [39]. Benders [40] showed that the master problem and the subproblem can be solved successively with information being communicated between them.

We put the vehicle selection decision variables into the master problem and put the distribution decision variables into the subproblem. We set an initial solution $\overline{y_{it,jt'}^k}$. The subproblem can be written as follows:

$$\begin{aligned} & Minimize \ \sum_{j \in N} \sum_{t \in T} (P_j^t - \sum_{k \in K} \sum_{t=0}^{t} x_{jt}^k) + \alpha \sum_{it,jt' \in A^T} \sum_{k \in K} (t' - t) \overline{y_{it,jt'}^k} \\ & s.t. \quad - \sum_{\forall jt' \in N^T} x_{jt'}^k \ge -Q_k \times \sum_{it,jt' \in A^T} y_{it,jt'}^k, \forall k \in K, \\ & \qquad - \sum_{k \in K} x_{jt}^k \ge -P_j^t, \forall t \in T, j \in N, \\ & \qquad x_{jt}^k \ge 0, integer, \forall jt \in N^T, k \in K. \end{aligned} \tag{12}$$

Therefore, the dual problem of Eq. 12, can be obtained below,

$$\begin{aligned} & Maximize \ \pi^T (-Q_k \sum_{it,jt' \in A^T} \overline{y_{it,jt'}^k} - P_j^t) \\ & s.t. \ \pi^T G \le f, \\ & \pi \ge 0. \end{aligned} \tag{13}$$

G denotes the transposed matrix of the constraint coefficient matrix in model (Eq. 12). f denotes the transposed matrix of objective function coefficient matrix in model (Eq. 12). π^T denotes the dual variables in model (Eq. 13).

There are three possible solution with respect to model (Eq. 13): (1) infeasible, then exit; (2) unbounded, in which case choose any unbounded extreme ray (denoted as $\overline{\pi^T}$)

and add a feasibility cut $\overline{\pi^T}(-Q_k \cdot \sum_{it,jt' \in A^T} \overline{y^k_{it,jt'}} - P^t_j) \le 0$ into the master problem; (3) bounded, in which case take an optimal solution (denoted as $\overline{\overline{\pi^T}}$) and add an optimality cut $\overline{\overline{\pi^T}}(-Q_k \cdot \sum_{it,jt' \in A^T} \overline{y^k_{it,jt'}} - P^t_j) \le \theta$ into the master problem.

Therefore, the master problem of the time-space network of the humanitarian logistics can be written as Eq. 14,

$$
\begin{aligned}
&Minimize\ \ \theta + \alpha \sum_{it,jt' \in A^T} \sum_{k \in K} (t' - t) y^k_{it,jt'} \\
&s.t. \sum_{it \in N^T} y^k_{it,jt'} = 1, \forall jt' \in N^T, k \in K, \\
&\sum_{jt' \in N^T} y^k_{it,jt'} = 1, \forall it \in N^T, k \in K, \\
&\sum_{ij \in A} y^k_{it,jt'} = 1, \forall t, t' \in T, k \in K, \\
&\sum_{it \in N^T} \sum_{k \in K} y^k_{it,jt'} = \sum_{it \in N^T} \sum_{k \in K} y^k_{jt',it}, \forall jt' \in N^T, \\
&\overline{\pi^T}(-Q_k \cdot \sum_{it,jt' \in A^T} y^k_{it,jt'} - P^t_j) \le 0, \\
&\overline{\overline{\pi^T}}\left(-Q_k \cdot \sum_{it,jt' \in A^T} y^k_{it,jt'} - P^t_j\right) \le \theta, \\
&y^k_{it,jt'} \in \{0,1\}, \forall it, jt' \in N^T, k \in K, \\
&\pi \in R.
\end{aligned}
\tag{14}
$$

Through the iteration to solve the master problem and the subproblem, we can obtain the optimal solution of the master problem $(\theta^*, {y^k_{it,jt'}}^*)$. We can use ${y^k_{it,jt'}}^*$ to solve the time-space network of the humanitarian logistics, we can obtain the optimal solution $({x^k_{jt}}^*, {y^k_{it,jt'}}^*)$.

5. Numerical tests

This section describes the computational results for the proposed mathematical model and solution algorithm. In Subsection 5.1, we describe a numerical example. In Subsection 5.2, we exhibit the computational results with a case design. In Subsection 5.3, we conduct the sensitivity analysis. All the tests in this section were tested on a Lenovo Y400 with Intel Core i5-3230M CPU, 2.60 GHz frequency and 4 GB memory.

5.1 A case study

We cite a case that can reflect the proposed model. We suppose that there are 8 nodes, 30 days, 4 vehicles. The maximum loading capacity of each vehicle is shown in Table 1. The marginal utility from node i to node j is shown in Table 2. S_j (0), E_j (0), I_j (0), and R_j (0) are shown in Table 3. Where $\beta = 0.0001$, $\sigma = 0.2$, $\gamma = 0.5$, $\lambda = 1$. To make the results of the total time cost and the total delay cost in the same magnitude, we set $\alpha = 20$.

Table 1 The maximum loading capacity of the vehicle k(kg)

Vehicle number	1	2	3	4
Vehicle capacity	458	468	574	542

Table 2 The travel time from node *i* to node *j*(day)

Node number	1	2	3	4	5	6	7	8
1	0	3	2	3	5	1	8	6
2	3	0	4	3	2	5	7	1
3	2	4	0	6	3	4	2	5
4	3	3	6	0	7	2	1	5
5	5	2	3	7	0	6	3	2
6	1	5	4	2	6	0	4	7
7	8	7	2	1	3	4	0	3
8	6	1	5	5	2	7	3	0

Table 3 The initial value of $S_j(0)$, $E_j(0)$, $I_j(0)$, and $R_j(0)$

	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$S_j(0)$	10746	11288	14204	11271	14071	11218	14646	11750
$E_j(0)$	0	0	0	0	0	0	0	0
$I_j(0)$	639	650	723	695	670	766	717	710
$R_j(0)$	0	0	0	0	0	0	0	0

5.2 Computational results

Fig. 3 shows that the variation trend of $S_j(t)$, $E_j(t)$, $I_j(t)$, and $R_j(t)$ in the whole rescue process. As we can see, $I_j(t)$ reduces first then increases, $I_j(t)$ is equal 0 at last. This result consistent with the mechanism of infectious diseases.

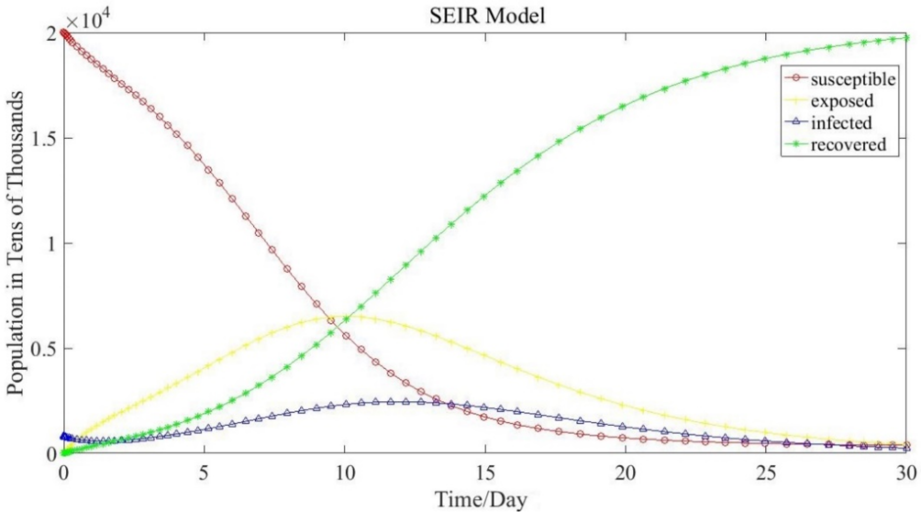


Fig. 3 The results of $S_j(t)$, $E_j(t)$, $I_j(t)$, and $R_j(t)$

Node 1 is distribution center. The driving route of each vehicle is node 1-6-4-7-3-5-2-8-1. Vehicles arrive time at node are 1, 3, 4, 6, 9, 11, 12, 18 days respectively. The delivery medicine resources number of all vehicles on each node are shown in Table 4. As shown in Table 4, the number of delivery medicine resource by all vehicles on node 6 is 143, 96, 172, and 208, respectively, and is bigger than other nodes. The medicine resources of other nodes are optimized according to the number and geographical location of infectious diseases. Therefore, this result conforms to reality according to the result of $S_j(t)$, $E_j(t)$, $I_j(t)$, and $R_j(t)$ in Fig. 3.

Fig. 4 shows the upper and lower bounds obtained by Benders decomposition algorithm are equal at the 12th iteration and will not change any more. Before that, the upper and lower bounds converge continuously. Experimental results show the effectiveness of the proposed model and algorithm.

Table 4 The delivery medicine resources number of all vehicles on each node

	node 1	node 2	node 3	node 4	node 5	node 6	node 7	node 8
Vehicle 1	0	98	33	65	76	143	20	23
Vehicle 2	0	20	107	47	54	96	83	61
Vehicle 3	0	45	131	23	74	172	48	81
Vehicle 4	0	71	60	40	8	208	75	90

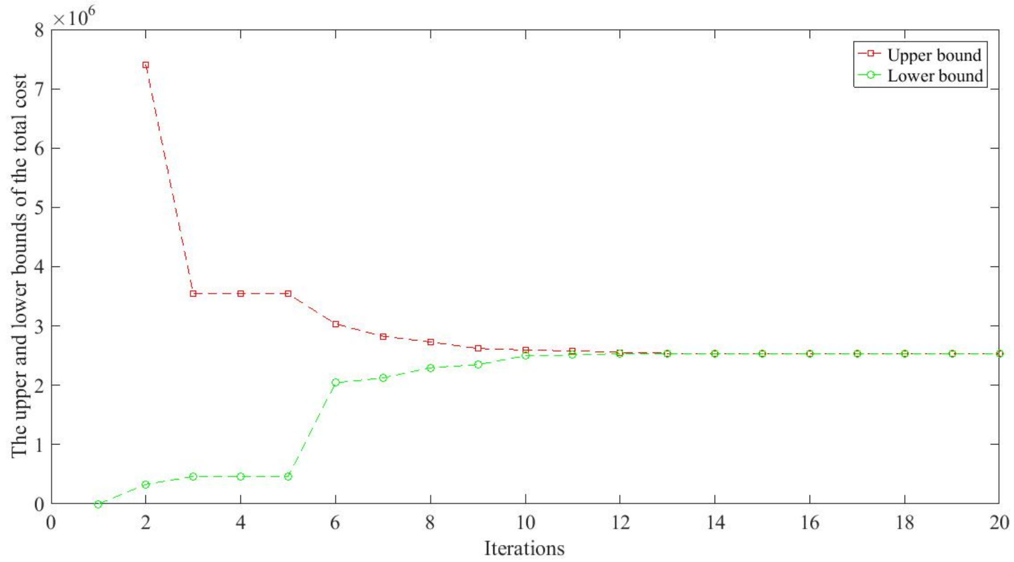


Fig. 4 Convergence of upper and lower bounds

5.3 Sensitivity analysis

In this section, a sensitivity analysis of the three key parameters (β, σ, γ) in the demand forecasting model is conducted. Fig. 5 shows the relationship between the different β and the number of infected people. Obviously, the greater β is, the more the number of infected people is, and the greater the gradient of demand is. Therefore, β should be selected appropriately in a practical problem.

As Fig. 6 shows, σ takes on five values ranging from 0.1 to 0.5. The larger σ is, the larger the demand is, and the greater the gradient of demand is. As Fig. 7 shows, γ also takes on five values ranging from 0.1 to 0.5. The lower σ is, the larger the demand is.

In this section, a sensitivity analysis of the three key parameters (β, σ, γ) in the demand forecasting model is conducted. Fig. 5 shows the relationship between the different β and the number of infected people. Obviously, the greater β is, the more the number of infected people is, and the greater the gradient of demand is. Therefore, β should be selected appropriately in a practical problem.

As Fig. 6 shows, σ takes on five values ranging from 0.1 to 0.5. The larger σ is, the larger the demand is, and the greater the gradient of demand is. As Fig. 7 shows, γ also takes on five values ranging from 0.1 to 0.5. The lower σ is, the larger the demand is.

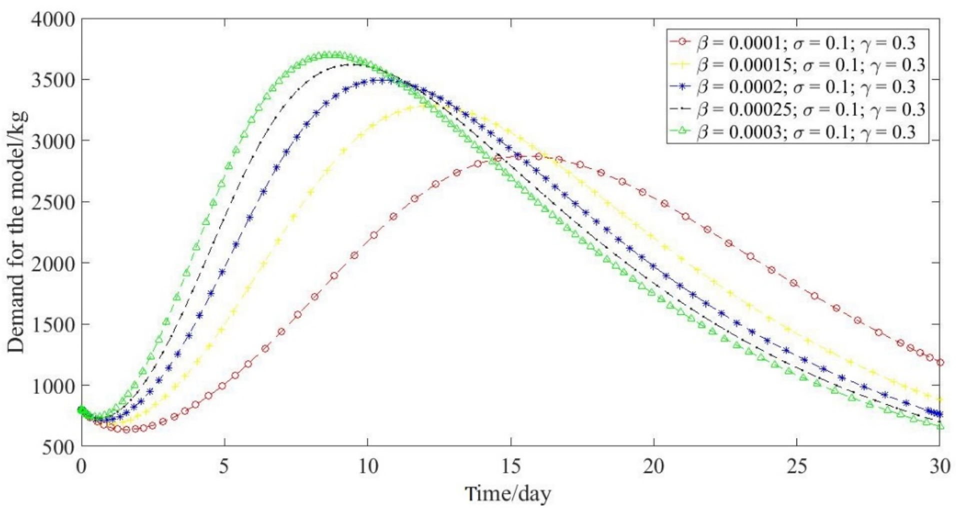


Fig. 5 The results of sensitivity analysis with different β

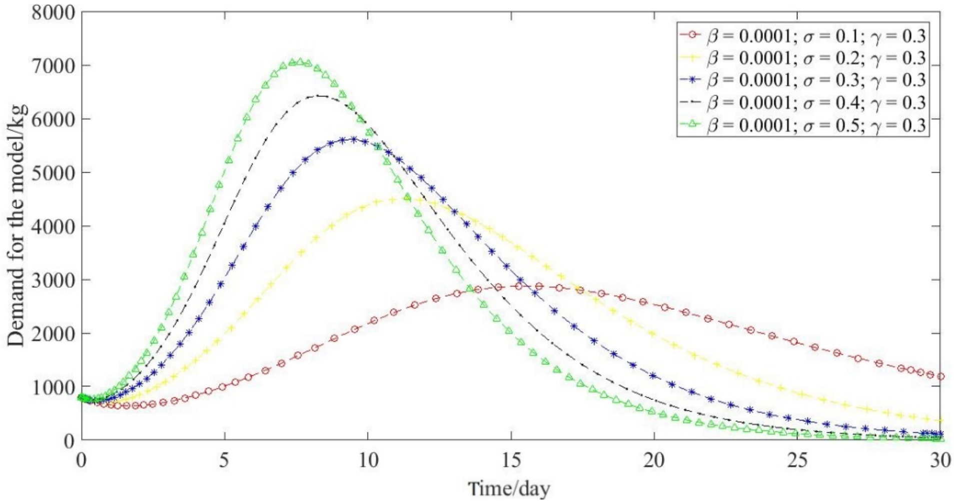


Fig. 6 The results of sensitivity analysis with different σ

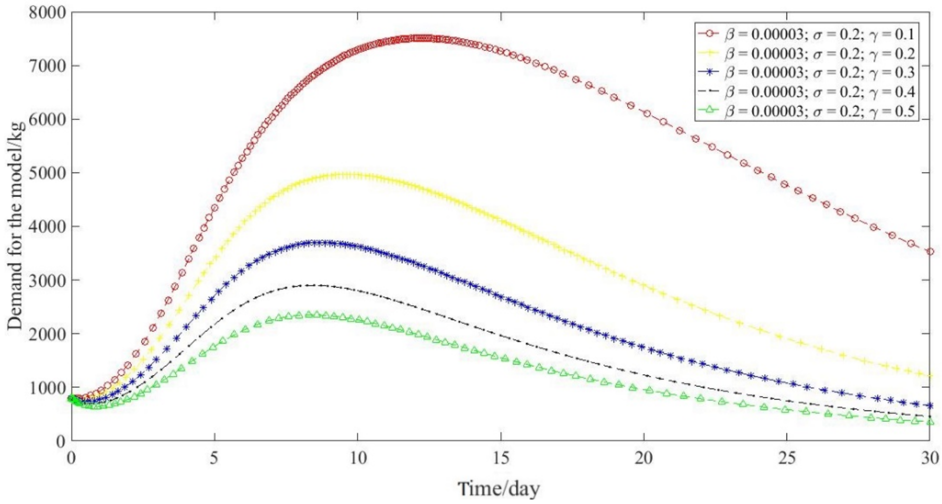


Fig. 7 The results of sensitivity analysis with different γ

5.4 Algorithm comparison

Genetic algorithm is a heuristic search and optimization technique inspired by natural evolution. They have been successfully applied to a wide range of real-world problems of significant complexity. In addition, ant colony algorithm is a class of metaheuristics which are inspired from the behaviour of real ants. The original idea consisted in simulating the stigmergic communication, therefore these algorithms are considered as a form of adaptive memory programming. Therefore, in order to compare the difference between Benders decomposition algorithm and genetic algorithm, ant colony algorithm, a time-space network model for the humanitarian logistics problem with logistics in controlling epidemic diffusion is also solved by a genetic algorithm and an ant colony algorithm in this paper.

Fig. 8 shows the optimization results of the genetic algorithm and the ant colony algorithm. The optimal value by the ant colony algorithm don't converge any more in 15th iteration, and the optimal value by the genetic algorithm don't converge any more in 23rd iteration. This result proves that the convergence performance and the computational accuracy of the ant colony algorithm is better than the genetic algorithm. The computing time of the Benders decomposition algorithm, the genetic algorithm and the ant colony algorithm is 10.8 s, 5.2 s and 8.9 s respectively. Therefore, the computational efficiency of the genetic algorithm is better than the other two algorithms. In addition, the optimal solution of the Benders decomposition algorithm, the genetic algorithm, and the ant colony algorithm is 2540084, 2854111, and 2540176, respectively. This result proves that the computational accuracy of the Benders decomposition algorithm is better than the genetic algorithm and the ant colony algorithm.

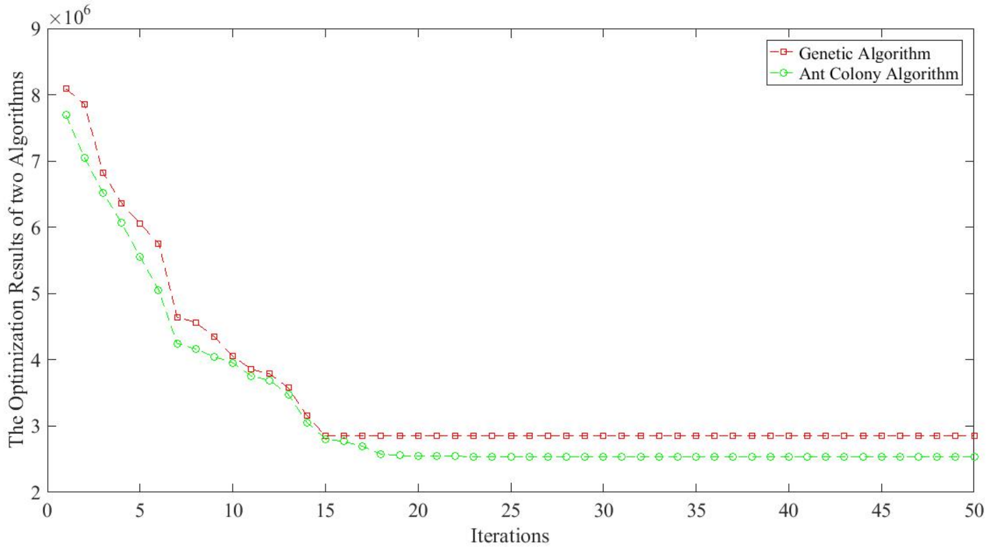


Fig. 8 The optimization results of the genetic algorithm and the ant colony algorithm

5.5 A larger case test

We expand the node and vehicle sizes based on the case in Subsection 5.1. We cite a case that can reflect the proposed model. We suppose that there are 16 nodes, 30 days, 8 vehicles. The maximum loading capacity of each vehicle, the marginal utility from node i to node j and $S_j(0)$, $E_j(0)$, $I_j(0)$, and $R_j(0)$ are expanded by the supposed node and vehicle sizes. Where β = 0.0001, σ = 0.2, γ = 0.5, λ = 1, α =20.

The optimization results of the larger case test by the Benders decomposition algorithm, the genetic algorithm and the ant colony algorithm are shown in Fig. 9. As Fig. 9 shows, the upper and lower bounds obtained by Benders decomposition algorithm are equal at the 14th iteration and will not change any more. Before that, the upper and lower bounds converge continuously.

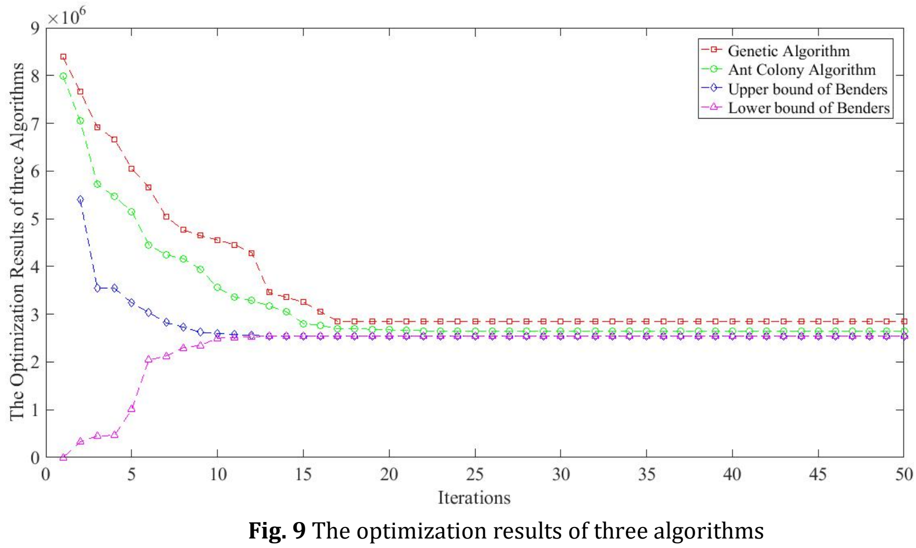


Fig. 9 The optimization results of three algorithms

The optimal value by the ant colony algorithm don't converge any more in 13th iteration, and the optimal value by the genetic algorithm don't converge any more in 24th iteration. Experimental results show the effectiveness of the proposed model and algorithm. The computing time of the Benders decomposition algorithm, the genetic algorithm and the ant colony algorithm is 40.7 s, 26.5 s and 33.4 s respectively. The computational efficiency of the genetic algorithm is better than the other two algorithms. In addition, the optimal solution of the Benders decomposition algorithm, the genetic algorithm and the ant colony algorithm is 2551213, 2896740, 2651176 respectively. This result proves that the computational accuracy of the Benders decomposition algorithm is better than the genetic algorithm and the ant colony algorithm. Therefore, the larger case experimental results show the validity of results in Subsections 5.2 and 5.4 and the effectiveness of the proposed model and algorithm.

6. Conclusion

In this paper, a time-space network model for the humanitarian logistics problem with logistics in controlling epidemic diffusion is proposed. We use the SEIR model which is a new differential equation to solve the dynamic information in humanitarian relief phase. A multistage mixed integer programming model for the humanitarian logistics and transport resource with a time-space network is proposed. The linear programming problem is solved by the Benders decomposition algorithm, the genetic algorithm and the ant colony algorithm. Finally, we use cases to calculate model and algorithm. The computational efficiency of the genetic algorithm is better than the other two algorithms. The results of cases prove that the computational accuracy of the Benders decomposition algorithm is better than the genetic algorithm and the ant colony algorithm and the correctness of the model and algorithm. This paper can improve the efficiency of emergency rescue and reduce the loss of people's lives and property caused by decision-making mistakes. Sensitivity analysis illustrate the effect of parameters on the result. Through sensitivity analysis, the influence of parameters on dynamic information is understood, and the internal relationship between route selection of emergency rescue vehicles and allocation of medical supplies is grasped, so as to provide reference for decision makers, and thus provide experience for future emergency decision-making. Due to the complexity of emergency rescue, only the time cost and delay cost in the process of emergency rescue are considered in this paper, so it is necessary to further study the comprehensiveness of emergency rescue.

Acknowledgement

This work is supported by grants from the National Social Science Foundation of China (No. 16ZDA054, Dr. Jie Cao) and Scientific and Technologial Innovation Programs of Higher Education Institutions in Shanxi (No. 2023L432, Dr. He Han).

References

[1] Wang, H., Wang, X., Zeng, A.Z. (2009). Optimal material distribution decisions based on epidemic diffusion rule and stochastic latent period for emergency rescue, *International Journal of Mathematics in Operational Research*, Vol. 1, No. 1-2, 76-96, doi: 10.1504/IJMOR.2009.022876.

[2] Huang, K., Jiang, Y., Yuan, Y., Zhao, L. (2015). Modeling multiple humanitarian objectives in emergency response to large-scale disasters, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 75, 1-17, doi: 10.1016/j.tre.2014.11.007.

[3] Cao, J., Han, H., Jiang, Y.-P., Wang, Y.-J. (2018). Emergency rescue vehicle dispatch planning using a hybrid algorithm, *International Journal of Information Technology & Decision Making*, Vol. 17, No. 6, 1865-1890, doi: 10.1142/S0219622018500414.

[4] Holguín-Veras, J., Pérez, N., Jaller, M., Van Wassenhove, L.N., Aros-Vera, F. (2013). On the appropriate objective function for post-disaster humanitarian logistics models, *Journal of Operations Management*, Vol. 31, No. 5, 262-280, doi: 10.1016/j.jom.2013.06.002.

[5] Zhang, J., Ma, Z. (2003). Global dynamics of an SEIR epidemic model with saturating contact rate, *Mathematical Biosciences*, Vol. 185, No. 1, 15-32, doi: 10.1016/S0025-5564(03)00087-7.

[6] Li, D., Liu, S., Cui, J. (2017). Threshold dynamics and ergodicity of an SIRS epidemic model with Markovian switching, *Journal of Differential Equations,* Vol. 263, No. 12, 8873-8915, doi: 10.1016/j.jde.2017.08.066.

[7] Fan, K., Zhang, Y., Gao, S., Wei, X. (2017). A class of stochastic delayed SIR epidemic models with generalized nonlinear incidence rate and temporary immunity, *Physica A: Statistical Mechanics and its Applications*, Vol. 481, 198-208, doi: 10.1016/j.physa.2017.04.055.

[8] Song, P., Lou, Y., Xiao, Y. (2019). A spatial SEIRS reaction-diffusion model in heterogeneous environment, *Journal of Differential Equations,* Vol. 267, No. 9, 5084-5114, doi: 10.1016/j.jde.2019.05.022.

[9] Yang, P., Wang, Y. (2019). Dynamics for an SEIRS epidemic model with time delay on a scale-free network, *Physica A: Statistical Mechanics and its Applications,* Vol. 527, Article No. 121290, doi: 10.1016/j.physa.2019.121290.

[10] Bolzoni, L., Bonacini, E., Soresina, C., Groppi, M. (2017). Time-optimal control strategies in SIR epidemic models, *Mathematical Biosciences*, Vol. 292, 86-96, doi: 10.1016/j.mbs.2017.07.011.

[11] Guo, W., Zhang, Q., Li, X., Wang, W. (2018). Dynamic behavior of a stochastic SIRS epidemic model with media coverage, *Mathematical Methods in the Applied Sciences*, Vol. 41, No. 14, 5506-5525, doi: 10.1002/mma.5094.

[12] Britton, T., Ouédraogo, D. (2018). SEIRS epidemics with disease fatalities in growing populations, *Mathematical Biosciences*, Vol. 296, 45-59, doi: 10.1016/j.mbs.2017.11.006.

[13] Anparasan, A., Lejeune, M. (2019). Resource deployment and donation allocation for epidemic outbreaks, *Annals of Operations Research*, Vol. 283, 9-32, doi: 10.1007/S10479-016-2392-0.

[14] Yee, W.S., Ng, H., Yap, T.T.V., Goh, V.T., Ng, K.H., Cher, D.T. (2022). An evaluation study on the predictive models of breast cancer risk factor classification, *Journal of Logistics, Informatics and Service Science,* Vol. 9, No. 3, 129-145, doi: 10.33168/LISS.2022.0310.

[15] Choi, J.-W., Jeong, E.-R. (2022). Multi-output convolutional neural network based distance and velocity estimation technique for orthogonal frequency division multiplexing radar systems, *Webology,* Vol. 19, No. 1, 4555-4570, doi: 10.14704/WEB/V19I1/WEB19302.

[16] Farahani, R.Z., Lotfi, M.M., Baghaian, A., Ruiz, R., Rezapour, S. (2020). Mass casualty management in disaster scene: A systematic review of OR&MS research in humanitarian operations, *European Journal of Operational Research,* Vol. 287, No. 3, 787-819, doi: 10.1016/j.ejor.2020.03.005.

[17] Kwapong Baffoe, B.O., Luo, W. (2020). Humanitarian relief sustainability: A framework of humanitarian logistics digital business ecosystem, *Transportation Research Procedia,* Vol. 48, 363-387, doi: 10.1016/j.trpro.2020.08.032.

[18] 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.

[19] Zaldívar-Colado, A., Tripp-Barba, C., Brito-Rojas, J.A., Aguilar-Calderón, J.A., García-Sánchez, O.V., Ramírez-Zambada, L., Zatarain-Montalvo, J.N. (2017). Management of traffic lights for emergency service, *Tehnički Vjesnik – Technical Gazette,* Vol. 24, No. 2, 643-648, doi: 10.17559/TV-20140910051747.

[20] Wang, L., Chen, X.Y., Zhang, H. (2021). Joint distribution models in fast-moving consumer goods wholesale enterprise: Comparative analysis and a case study, *Advances in Production Engineering & Management*, Vol. 16, No. 2, 212-222, doi: 10.14743/apem2021.2.395.

[21] Zhou, Y., Liu, J., Zhang, Y., Gan, X. (2017). A multi-objective evolutionary algorithm for multi-period dynamic emergency resource scheduling problems, *Transportation Research Part E: Logistics and Transportation Review,* Vol. 99, 77-95, doi: 10.1016/j.tre.2016.12.011.

[22] Othman, S.B., Zgaya, H., Dotoli, M., Hammadi, S. (2017). An agent-based decision support system for resources' scheduling in emergency supply chains, *Control Engineering Practice*, Vol. 59, 27-43, doi: 10.1016/j.conengprac.2016.11.014.

[23] Ferrer, J.M., Martín-Campo, F.J., Ortuño, M.T., Pedraza-Martínez, A.J., Tirado, G., Vitoriano, B. (2018). Multi-criteria optimization for last mile distribution of disaster relief aid: Test cases and applications, *European Journal of Operational Research*, Vol. 269, No. 2, 501-515, doi: 10.1016/j.ejor.2018.02.043.

[24] Garrido, R.A., Aguirre, I. (2020). Emergency logistics for disaster management under spatio-temporal demand correlation: The earthquakes case, *Journal of Industrial and Management Optimization*, Vol. 16, No. 5, 2369-2387, doi: 10.3934/jimo.2019058.

[25] Lam, E., Van Hentenryck, P., Kilby, P. (2020). Joint vehicle and crew routing and scheduling, *Transportation Science*, Vol. 54, No. 2, 488-511, doi: 10.1287/trsc.2019.0907.

[26] Rajak, S., Parthiban, P., Dhanalakshmi, R. (2019). A hybrid metaheuristics approach for a multi-depot vehicle routing problem with simultaneous deliveries and pickups, *International Journal of Mathematics in Operational Research*, Vol. 15, No. 2, 197-210, doi: 10.1504/IJMOR.2019.101619.

[27] Sedehzadeh, S., Seifbarghy, M., (2021). Redesigning a closed loop food supply chain network considering sustainability and food banks with different returns, *Economic Computation And Economic Cybernetics Studies and Research*, Vol. 55, No. 4, 85-100, doi: 10.24818/18423264/55.4.21.06.

[28] Gutjahr, W.J., Nolz, P.C. (2016). Multicriteria optimization in humanitarian aid, *European Journal of Operational Research*, Vol. 252, No. 2, 351-366, doi: 10.1016/j.ejor.2015.12.035.

[29] Liu, M., Xiao, Y. (2015). Optimal scheduling of logistical support for medical resource with demand information updating, *Mathematical Problems in Engineering*, Vol. 2015, Article ID 765098, doi: 10.1155/2015/765098.

[30] Liu, M., Zhang, Z., Zhang, D. (2015). A dynamic allocation model for medical resources in the control of influenza diffusion, *Journal of Systems Science and Systems Engineering,* Vol. 24, No. 3, 276-292, doi: 10.1007/s11518-015-5276-y.

[31] Buschiazzo, M., Mula, J., Campuzano-Bolarin, F. (2020). Simulation optimization for the inventory management of healthcare supplies, *International Journal of Simulation Modelling*, Vol. 19, No. 2, 255-266, doi: 10.2507/IJSIMM19-2-514.

[32] Fischetti, M., Ljubić, I., Sinnl, M. (2017). Redesigning benders decomposition for large-scale facility location, *Management Science*, Vol. 63, No. 7, 2146-2162, doi: 10.1287/MNSC.2016.2461.

[33] Alkaabneh, F., Diabat, A., Gao, H.O. (2020). Benders decomposition for the inventory vehicle routing problem with perishable products and environmental costs, *Computers & Operations Research,* Vol. 113, Article No. 104751, doi: 10.1016/j.cor.2019.07.009.

[34] Fachini, R.F., Armentano, V.A. (2020). Logic-based Benders decomposition for the heterogeneous fixed fleet vehicle routing problem with time windows, *Computers & Industrial Engineering*, Vol. 148, Article No. 106641, doi: 10.1016/j.cie.2020.106641.

[35] Cordeau, J.-F., Furini, F., Ljubić, I. (2019). Benders decomposition for very large scale partial set covering and maximal covering location problems, *European Journal of Operational Research*, Vol. 275, No. 3, 882-896, doi: 10.1016/j.ejor.2018.12.021.

[36] Behmanesh, R., Rahimi, I. (2021). Improved ant colony optimization for multi-resource job shop scheduling: A special case of transportation, *Economic Computation And Economic Cybernetics Studies And Research,* Vol. 55, No. 4, 277-294, doi: 10.24818/18423264/55.4.21.18.

[37] Tang, F. (2021) An improved intelligent bionic optimization algorithm based on the growth characteristics of tree branches, *Journal of Intelligent & Fuzzy Systems*, Vol. 40, No. 3, 3821-3829, doi: 10.3233/JIFS-190487.

[38] Liu, M., Liang, J. (2013). Dynamic optimization model for allocating medical resources in epidemic controlling, *Journal of Industrial Engineering and Management,* Vol. 6, No. 1, 73-88, doi: 10.3926/jiem.663.

[39] Han, H., Cao, J., Wang, Y.-J. (2023). A hybrid algorithm of ant colony and benders decomposition for large-scale mix-integer linear programming, *International Journal of Information Technology & Decision Making*, Vol. 22, 1-23, doi: 10.1142/S0219622023500335.

[40] Benders, J.F. (1962). Partitioning procedures for solving mixed-variables programming problems, *Numerische Mathematik,* Vol. 4, 238-252, doi: 10.1007/BF01386316.

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 3 | September 2023 | pp 267-398

Contents

Published by CPE, University of Maribor

ISSN 1854-6250

apem-journal.org