ISSN 1854-6250

APEM journal

Advances in Production Engineering & Management

Volume 13 | Number 4 | December 2018

University of Maribor

Published by CPE
apem-journal.org

Advances in Production Engineering & Management

Identification Statement

APEM journal

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

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

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

University of Maribor

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

APEM Editorial

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

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

University of Maribor
Chair of Production Engineering (CPE)

Advances in Production Engineering & Management

Volume 13 | Number 4 | December 2018 | pp 369–506

Contents

Scope and topics **372**

Flexible job shop scheduling using zero-suppressed binary decision diagrams **373**
Meolic, R.; Brezočnik, Z.

Functional objectives decision-making of discrete manufacturing system based on integrated ant colony optimization and particle swarm optimization approach **389**
Xu, W.; Yin, Y.

Effect of process parameters on the surface roughness of aluminum alloy AA 6061-T6 sheets in frictional stir incremental forming **405**
Azpen, Q.; Baharudin, H.; Sulaiman, S.; Mustapha, F.

Visual and optometric issues with smart glasses in Industry 4.0 working environment **417**
Vujica Herzog, N.; Buchmeister, B.; Beharic, A.; Gajsek, B.

Multi-objective production planning model for equipment manufacturing enterprises with multiple uncertainties in demand **429**
Liu, Y.F.; Zhang, Q.S.

A quantitative analysis method of greenhouse gas emission for mechanical product remanufacturing based on Petri net **442**
Shi, J.L.; Fan, S.J.; Wang, Y.J.; Cheng, J.S.

Comprehensive analysis and study of the machinability of a high strength aluminum alloy (EN AW-AlZn5.5MgCu) in the high-feed milling **455**
Duplák, J.; Hatala, M.; Dupláková, D.; Steranka, J.

Hybrid fruit fly optimization algorithm for solving multi-compartment vehicle routing problem in intelligent logistics **466**
Wang, C.L.; Li, S.W.

Multi-objective transport network design with a reversible simulated annealing algorithm **479**
Feng, X.; Ruan, Z.; Zhu, X.; Zhang, L.

Compatibility of ionic liquids with hydraulic system components **492**
Kambič, M.; Kalb, R.; Tič, V.; Lovrec, D.

Calendar of events **504**

Notes for contributors **505**

Journal homepage: apem-journal.org

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

Scope and topics

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

Fields of interest include, but are not limited to:

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

APEM journal
Advances in Production Engineering & Management
Volume 13 | Number 4 | December 2018 | pp 373–388
https://doi.org/10.14743/apem2018.4.297
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Flexible job shop scheduling using zero-suppressed binary decision diagrams

Meolic, R.[a], **Brezočnik, Z.**[a,*]

[a]University of Maribor, Faculty of Electrical Engineering and Computer Science, Maribor, Slovenia

ABSTRACT

A flexible job shop scheduling problem (FJSP) is a widely studied NP-hard combinatorial problem. Its goal is to optimise the production plans for simultaneously produced parts, where each part production consists of executing various operations. Each operation can be executed on several, or even all, available machines. A distinctive subproblem of FJSP is the identification of feasible solutions. A feasible solution is an allocation plan (i.e. assignment of a machine to a particular operation of a part to be produced) yielding an execution schedule satisfying the given resource constraints. FJSP is applied primarily in manufacturing systems, but it can be used to optimise Internet traffic, cloud computing, and other resource scheduling problems as well. So far, the exact methods for solving FJSP have not been considered attractive, since they seemed incapable of coping with real-size problems. This paper proposes a novel exact approach to solving FJSP which can find and count out all schedules of relatively large systems. The approach is successful due to the power of a special data structure called zero-suppressed binary decision diagrams to represent and manipulate the set of all feasible solutions efficiently. All the algorithms are implemented and tested by using our free Binary Decision Diagram package called Biddy.

ARTICLE INFO

Keywords:
Process planning;
Exact optimization;
Flexible job shop scheduling;
Unate cube set algebra;
Zero-suppressed binary decision diagram

**Corresponding author:*
zmago.brezocnik@um.si
(Brezočnik, Z.)

Article history:
Received 17 June 2018
Revised 13 November 2018
Accepted 15 November 2018

1. Introduction

Manufacturing sites have to plan their production to ensure the profitability of manufacturing products, resource utilization, and product delivery time. Production planning involves many attributes that can be categorised into different domains [1, 2]. Process planning and scheduling are the two most essential tasks in a manufacturing company [3]. From the computer science perspective, they are formulated together as a job shop scheduling problem. The job shop scheduling problem is about the optimization of the production of several parts which are produced simultaneously. Each part can be produced by one or more sequences of operations. Operations are executed using a given set of machines. In a flexible job shop scheduling problem (FJSP), all machines can perform all operations (total flexibility), or each machine can execute a subset of operations (partial flexibility). The processing time for operations varies on different machines. The following assumptions are also very common: parts are produced independently of each other, sequences of operations and processing times are fixed and known in advance, setup time and transport time are either negligible, or included in the processing time, all machines are available all the time, the operation execution cannot be interrupted.

Methods for FJSP can be categorised either as exact or approximation. Exact methods can obtain an exact optimal solution. However, they do not scale well for solving large FJSP prob-

lems. Therefore, most of the approaches to FJSP resort to approximation methods such as genetic and evolutionary algorithms. Such methods can solve large-scale problems, but may lack either local or global search ability. Our method of process planning and scheduling is oriented towards the generation of the exact solution, and is based on the efficient realization of unate cube set algebra with Zero-suppressed Binary Decision Diagrams (ZBDDs) [4, 5]. The previous work that has influenced the proposed method the most are [1] and [6]. Takahashi *et al.* [1] used ZBDDs to represent solution candidates that satisfy the constraints in process planning. Jensen *et al.* [6] implemented the task graph scheduling problem with uniform processors and arbitrary task execution times as a state space exploration problem, and solved it with Binary Decision Diagrams (BDDs) [7], but they did not use the unate cube set algebra and ZBDDs.

The remainder of this paper has the following structure. The framework for FJSP is specified in Section 2. Generating of feasible solutions is discussed in Section 3. The novel scheduling algorithm is introduced in Section 4. Section 5 gives the necessary background on ZBDDs, operations in unate cube set algebra, and gives complete algorithms for feasible solutions' generation and scheduling. Results of experimental studies are reported in Section 6. Section 7 describes the conclusions and challenges of future work.

2. Specification of a FJSP test case

For specifying the framework for FJSP we use the approach from [1]. The factory production programme that will be observed in this paper is given in Fig. 1. It contains eight different parts, denoted as P_s, $s = 1, \ldots, 8$. Each part can be produced in several different ways, called *process sequences.* An operation type O_w is performed at each step of a process sequence. The number of operations is 15. The order of operations within a process sequence is important. Operations are denoted with squared nodes, and the precedence relation between the nodes with an arrow. Moreover, the graphical representation includes two types of brackets. The vertical bars »||« enclosing nodes indicate alternative sequences of the enclosed nodes, and the square brackets »[]« mean the arbitrary order of the bracketed nodes. For example, in the manufacturing of P_1, there are three selectable alternative process sequences: $O_1 \rightarrow O_2 \rightarrow O_3$, $O_3 \rightarrow O_5 \rightarrow O_4$, or $O_3 \rightarrow O_1 \rightarrow O_4$. On the other hand, for example, any process sequence in P_3 begins with O_4 followed by O_7 or with O_7 followed by O_4. Thus, the process sequences for P_3 are $O_4 \rightarrow O_7 \rightarrow O_8$, $O_4 \rightarrow O_7 \rightarrow O_9$, $O_7 \rightarrow O_4 \rightarrow O_8$, or $O_7 \rightarrow O_4 \rightarrow O_9$. Finding process sequences for other parts is straightforward.

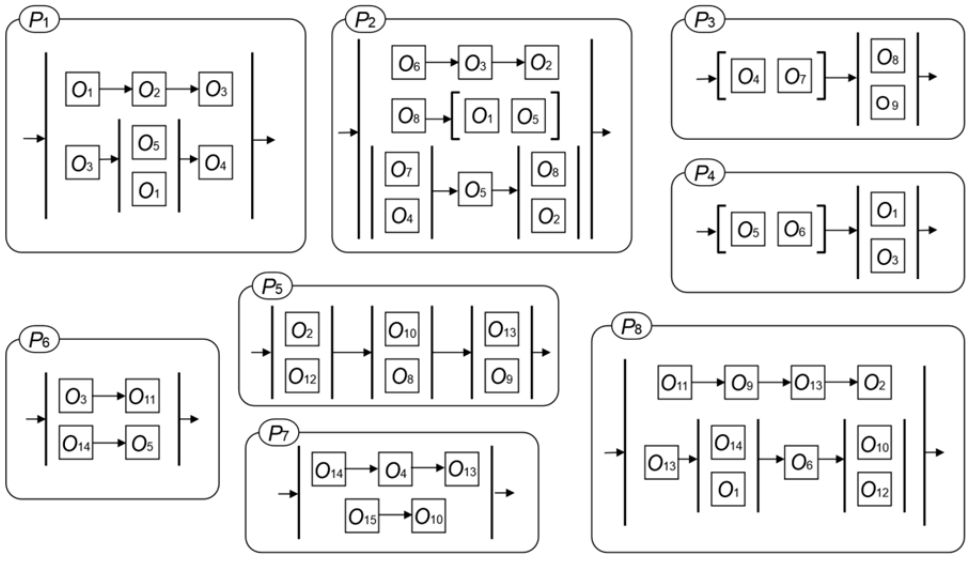


Fig. 1 Process sequences for a factory producing eight parts

In the factory, the production of parts is performed by a set of machine instances of different machine types $M_i, i = 1, \dots, 8$. Each machine type has a given number of instances with the same performance. Table 1 shows information on the available machine types, operations, number of machine instances of each machine type, and the processing time for operation O_w on machine type M_i. For example, O_1can be processed on an instance of the machine type M_1, M_2, or M_7. The processing time of O_1 on the machine type M_1, M_2, or M_7 is 6, 5, or 3 time units, respectively. All the values in Table 1 are taken from [2].

A set of machine instances installed in a factory is called the *factory configuration*. We assume that the maximum number of installable machine instances in the factory is equal to or less than the *factor capacity*. For example, if the factory capacity is 3, then the set of valid factory configurations includes the following ones: one instance of M_1, M_2, and M_3, two instances of M_1 and one instance of M_2 and also one instance of M_2 and one instance of M_3 as well, because the number of installed machines can be smaller than the factory capacity.

Table 1 Information on operations, machine types, number of machine type instances, and processing time

	Operations														
Machine type (#instances)	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9	O_{10}	O_{11}	O_{12}	O_{13}	O_{14}	O_{15}
M_1 (3)	6	-	-	6	9	-	2	-	-	-	-	-	3	-	-
M_2 (1)	5	8	-	-	-	5	4	-	-	-	-	4	-	-	-
M_3 (1)	-	-	10	-	8	-	-	6	-	-	8	-	-	-	-
M_4 (2)	-	-	8	4	-	6	-	-	3	-	-	-	-	10	-
M_5 (1)	-	-	-	-	7	5	-	-	-	-	5	-	-	-	8
M_6 (1)	-	4	-	-	-	-	1	-	-	5	-	2	2	-	-
M_7 (2)	3	-	-	-	-	-	-	-	-	2	-	1	-	-	12
M_8 (1)	-	-	-	-	-	-	-	-	5	-	3	-	4	14	-

3. Generating feasible solutions for process and resource planning with limits

A feasible solution is a combination of process sequences, machine types, and valid factory configuration which can be used to produce all required parts. For example, if a factory produces only part P_1 and factory capacity is 3, then one of the feasible solutions is a process sequence $O_3 \rightarrow O_1 \rightarrow O_4$, combined with the information that two instances of machine type M_1 and one instance of machine type M_3 are installed, and operation O_3 is processed on the machine type M_3, operation O_1 is processed on the first instance of the machine type M_1, and operation O_4 is processed on the second instance of the machine type M_1. Another feasible solution is a similar one, where both O_1 and O_4 are processed in the first instance of the machine type M_1. It is expected that the first mentioned feasible solution would be preferred, due to the fact that the part will be completed earlier, since operations O_1 and O_4 can be carried out in parallel.

3.1 Encoding the process planning problem with unate cube set algebra

Unate cube set algebra is a mathematical theory about manipulation of combination sets [4, 5]. For the given *universal set* of elements, a *combination set* (or a *cube set*) is a set of its subsets also called *cubes*. For example, if the universal set is {a,b,c,d}, then some of the possible combination sets are {}, {{}}, {{a},{b,d}}, and {{},{b,c},{b,c,d}}. The first one from this list is an *empty set*. The second one, which consists of the *empty cube* only, is called a *base set* (in some publications it is called a *unit set*). We shorten the notation of combination sets such that {{},{a},{b,c},{b,c,d}} is written as {1;a;bc;bcd}.

In encoding the process planning problem, the elements of a universal set are as follows:

- one $O^s_{w,r}$ element (called the O-variable) for each possible combination of values s, w and r denoting operation O_w, performed on the r-th place of a process sequence for part P_s.
- one $MX_{i,j}$ element (called the MX-variable) for each possible combination of values i and j, denoting an instance j for machine type M_i.

- one $M_i^{s,w,r}$ element (called the M-variable) for each possible combination of values s, w, r and i, denoting a machine type M_i for operation O_w, performed on the r-th place of a process sequence for part P_s.

For compatibility with [1] and [2], from now on, we write MX-variables without the letter X. Thus, M-variables and MX-variables differ only in the number of indices; the former ones have 4 and the latter 2 indices, respectively.

3.2 Algorithms for process and resource planning

There are several algorithms involved in the calculation of feasible solutions for process and resource planning, which are applied consecutively. To illustrate the ideas of the algorithms involved, we use a manageable small example, consisting only of part P_3 from Fig. 1 and machine types M_1, M_2, and M_3 from Table 1. For the sake of simplicity, let us set the factory capacity to 3.

The calculation starts by generating all possible *process sequences* for the production of P_3. Process sequences are encoded with cubes. According to the semantics of the graphical representation explained in Subsection 2.1, part P_3 is described with a set X^3 consisting of four cubes:

$$X^3 = \{O^3_{4,1}O^3_{7,2}O^3_{8,3}; O^3_{4,1}O^3_{7,2}O^3_{9,3}; O^3_{7,1}O^3_{4,2}O^3_{8,3}; O^3_{7,1}O^3_{4,2}O^3_{9,3}\} \tag{1}$$

In the second step, a process plan is generated for each part. A process plan is a set of all possible process sequences, extended with the information about the suitable machine types for processing each operation. Because there are many possibilities to complete an operation (e.g., operation O_7 can be completed either on machine type M_1 or M_2), we get more combinations than after the initial step. For our simple example, we get four combinations:

$$X^3 = \begin{Bmatrix} O^3_{4,1}M^{3,4,1}_1 O^3_{7,2}M^{3,7,2}_1 O^3_{8,3}M^{3,8,3}_3; O^3_{4,1}M^{3,4,1}_1 O^3_{7,2}M^{3,7,2}_2 O^3_{8,3}M^{3,8,3}_3; \\ O^3_{7,1}M^{3,7,1}_1 O^3_{4,2}M^{3,4,2}_1 O^3_{8,3}M^{3,8,3}_3; O^3_{7,1}M^{3,7,1}_2 O^3_{4,2}M^{3,4,2}_1 O^3_{8,3}M^{3,8,3}_3 \end{Bmatrix} \tag{2}$$

Please note that by assigning machine types to the process sequences, we also cut sequences that are not realisable. For example, operation O_9 is not used in any cube because it cannot be completed using the available set of machine types.

The third step is the construction of a *comprehensive process plan X* by computing a Cartesian product of all parts' process plans under the assumption that parts can be produced in parallel. They are not completely independent of each other, because they share the available machine types. However, in our small example, which illustrates the step-by-step evolving of the set of feasible solutions, the factory produces only one part, therefore, the comprehensive process plan is the same as the obtained process plan in the previous step ($X = X^3$). In general, the number of cubes in the comprehensive process plan is a product of cube numbers of all the produced parts.

In the fourth step, a set of MX-variables is added to each cube. The resulting set of feasible solutions is called a *process and resource plan*. All possible factory configurations should be considered up to three installed machine type instances. A process and resource plan is a huge set for any non-trivial problem. For our small example, we get a set of 12 cubes:

$$X = \begin{Bmatrix} O^3_{4,1}M^{3,4,1}_1 O^3_{7,2}M^{3,7,2}_1 O^3_{8,3}M^{3,8,3}_3 M_{1,1}M_{3,1}M_{1,2}M_{1,3}; O^3_{4,1}M^{3,4,1}_1 O^3_{7,2}M^{3,7,2}_1 O^3_{8,3}M^{3,8,3}_3 M_{1,1}M_{3,1}M_{1,2}; \\ O^3_{4,1}M^{3,4,1}_1 O^3_{7,2}M^{3,7,2}_1 O^3_{8,3}M^{3,8,3}_3 M_{1,1}M_{3,1}; O^3_{4,1}M^{3,4,1}_1 O^3_{7,2}M^{3,7,2}_2 O^3_{8,3}M^{3,8,3}_3 M_{1,1}M_{2,1}M_{3,1}M_{1,2}M_{1,3}; \\ O^3_{4,1}M^{3,4,1}_1 O^3_{7,2}M^{3,7,2}_2 O^3_{8,3}M^{3,8,3}_3 M_{1,1}M_{2,1}M_{3,1}M_{1,2}; O^3_{4,1}M^{3,4,1}_1 O^3_{7,2}M^{3,7,2}_2 O^3_{8,3}M^{3,8,3}_3 M_{1,1}M_{2,1}M_{3,1}; \\ O^3_{7,1}M^{3,7,1}_1 O^3_{4,2}M^{3,4,2}_1 O^3_{8,3}M^{3,8,3}_3 M_{1,1}M_{3,1}M_{1,2}M_{1,3}; O^3_{7,1}M^{3,7,1}_1 O^3_{4,2}M^{3,4,2}_1 O^3_{8,3}M^{3,8,3}_3 M_{1,1}M_{3,1}M_{1,2}; \\ O^3_{7,1}M^{3,7,1}_1 O^3_{4,2}M^{3,4,2}_1 O^3_{8,3}M^{3,8,3}_3 M_{1,1}M_{3,1}; O^3_{7,1}M^{3,7,1}_2 O^3_{4,2}M^{3,4,2}_1 O^3_{8,3}M^{3,8,3}_3 M_{1,1}M_{2,1}M_{3,1}M_{1,2}M_{1,3}; \\ O^3_{7,1}M^{3,7,1}_2 O^3_{4,2}M^{3,4,2}_1 O^3_{8,3}M^{3,8,3}_3 M_{1,1}M_{2,1}M_{3,1}M_{1,2}; O^3_{7,1}M^{3,7,1}_2 O^3_{4,2}M^{3,4,2}_1 O^3_{8,3}M^{3,8,3}_3 M_{1,1}M_{2,1}M_{3,1} \end{Bmatrix} \tag{3}$$

In the presented result we have taken into account the fact that configurations of the same length consisting of the same machine types (e.g. $\{M_{1,1}, M_{1,2}\}, \{M_{1,1}, M_{1,3}\}$, and $\{M_{1,2}, M_{1,3}\}$), yield symmetric results and, thus, it is enough to keep only one of them. We keep a configuration

consisting of variables that are declared earlier (in this case $\{M_{1,1}, M_{1,2}\}$), and remove the others. A less effective approach is used in [1] and [2]. There, the authors first generate a process and resource plan without considering the mentioned redundancy, and then have to restrict the result by an extra algorithm.

In the last step, the result is limited to the factory capacity. Only six cubes meet the requirement that the maximum number of the installed machine types instances in the factory is three. This last step gives us the set of all feasible solutions:

$$X = \left\{ \begin{array}{c} O_{4,1}^{3} M_1^{3,4,1} O_{7,2}^{3} M_1^{3,7,2} O_{8,3}^{3} M_3^{3,8,3} M_{1,1} M_{3,1} M_{1,2};\ O_{4,1}^{3} M_1^{3,4,1} O_{7,2}^{3} M_1^{3,7,2} O_{8,3}^{3} M_3^{3,8,3} M_{1,1} M_{3,1}; \\ O_{4,1}^{3} M_1^{3,4,1} O_{7,2}^{3} M_2^{3,7,2} O_{8,3}^{3} M_3^{3,8,3} M_{1,1} M_{2,1} M_{3,1};\ O_{7,1}^{3} M_1^{3,7,1} O_{4,2}^{3} M_1^{3,4,2} O_{8,3}^{3} M_3^{3,8,3} M_{1,1} M_{3,1} M_{1,2}; \\ O_{7,1}^{3} M_1^{3,7,1} O_{4,2}^{3} M_1^{3,4,2} O_{8,3}^{3} M_3^{3,8,3} M_{1,1} M_{3,1};\ O_{7,1}^{3} M_2^{3,7,1} O_{4,2}^{3} M_1^{3,4,2} O_{8,3}^{3} M_3^{3,8,3} M_{1,1} M_{2,1} M_{3,1} \end{array} \right\} \quad (4)$$

4. Job shop scheduling algorithm

In general, the job shop scheduling problem consists of two subproblems: determining the best feasible solution, and determining the best schedule for this solution. After creating the set of all feasible solutions, we solve these subproblems jointly. The state space is examined step by step until we find the first set of feasible solutions (containing either one single solution or more solutions) that completes the production of all parts. Alternatively, the computation may continue until all feasible solutions are examined.

The straightforward criterion for job shop scheduling is the amount of time, called *makespan*, needed to produce all parts. Other criteria can be observed as well, such as the *total workload* (the sum of working times of all machines), and the *workload of the critical machine* (the working time of the most loaded machine).

4.1 Extending the universal set

The proposed scheduling algorithm is an extension of the approach to generate feasible solutions. A similar scheduling algorithm was introduced in [6], but it is not based on unate cube set algebra.

For scheduling, we extend the universal set with the following elements:

- one $W^{s,r}$ element (called the W-variable) for each possible combination of values s and r, denoting that processing of the part type P_s is waiting to start the r-th operation,
- one $R^{s,r}$ element (called the R-variable) for each possible combination of values s and r, denoting that production of the part type P_s is running the r-th operation,
- one $B_{i,j}$ element (called the B-variable) for each possible combination of values i and j, denoting that the j-th instance of machine type M_i is busy,
- several T_n^s elements (called T-variables) for each possible combination of values s and n, denoting that processing of the part type P_s needs n time units to complete,
- one $S_{i,j}^{s,r}$ element (called the S-variable) for each possible combination of values s, r, i, and j, denoting that the r-th operation of a process sequence of the part type P_s is scheduled to a j-th instance of machine type M_i,
- one FS_n element (called the FS variable) for each possible value of n, denoting that n different machines type instances (not n different machines types) are needed, and
- some $G_t^{s,r}$ elements (called G variables), denoting that the r-th operation of a process sequence of the part type P_s started after t time units from the beginning of the scheduling.

4.2 Scheduling algorithm

Let us continue the example from Subsection 3.2, where six feasible solutions were found. Before the scheduling begins, the initialization is carried out. It transforms the set of feasible solutions into a *working set of cubes* in the following way:

1. The size of factory configuration is added to each cube by appending the appropriate FS-variable. These variables help sorting the obtained solutions at the end, and also make scheduling algorithm more efficient.
2. Each part is marked with "waiting to start the first operation" by adding the appropriate W-variable.

In our small example that illustrates the production of a single part P_3, all cubes require a factory with either two or three machine instances. The working set of cubes X is obtained immediately after the initialization by transforming the set of feasible solution (Eq. 4):

$$X = \begin{Bmatrix} W^{3,1}O^{3}_{4,1}M^{3,4,1}_{1}O^{3}_{7,2}M^{3,7,2}_{1}O^{3}_{8,3}M^{3,8,3}_{3}M_{1,1}M_{3,1}M_{1,2}FS_3; \\ W^{3,1}O^{3}_{4,1}M^{3,4,1}_{1}O^{3}_{7,2}M^{3,7,2}_{1}O^{3}_{8,3}M^{3,8,3}_{3}M_{1,1}M_{3,1}FS_2; \\ W^{3,1}O^{3}_{4,1}M^{3,4,1}_{1}O^{3}_{7,2}M^{3,7,2}_{2}O^{3}_{8,3}M^{3,8,3}_{3}M_{1,1}M_{2,1}M_{3,1}FS_3; \\ W^{3,1}O^{3}_{7,1}M^{3,7,1}_{1}O^{3}_{4,2}M^{3,4,2}_{1}O^{3}_{8,3}M^{3,8,3}_{3}M_{1,1}M_{3,1}M_{1,2}FS_3; \\ W^{3,1}O^{3}_{7,1}M^{3,7,1}_{1}O^{3}_{4,2}M^{3,4,2}_{1}O^{3}_{8,3}M^{3,8,3}_{3}M_{1,1}M_{3,1}FS_2; \\ W^{3,1}O^{3}_{7,1}M^{3,7,1}_{2}O^{3}_{4,2}M^{3,4,2}_{1}O^{3}_{8,3}M^{3,8,3}_{3}M_{1,1}M_{2,1}M_{3,1}FS_3 \end{Bmatrix} \quad (5)$$

Scheduling consists of three phases. In the first phase, every cube in the working set is checked whether it includes a part waiting to start the next operation. If such a part exists, then it is scheduled for all appropriate free machine instances. Indeed, if there is more than one such instance, the number of cubes in the working set is increased. All of the waiting parts are scheduled simultaneously. All parts that can be scheduled are scheduled (see Subsection 5.4 for notes about omitting this requirement). In any case, two parts cannot both be scheduled to the same machine instance. More formally, the transformation is described by Rule TR1.

Rule TR1: If a cube includes $W^{s,r}$, $M_i^{s,w,r}$, $M_{i,j}$, and does not include $R^{s,r}$ and $B_{i,j}$, then remove $M_i^{s,w,r}$ and $W^{s,r}$ and add $R^{s,r}$, $B_{i,j}$, $S_{i,j}^{s,r}$, T_n^s, and $G_t^{s,r}$, where n is the number of time units needed to complete operation O_w on machine type M_i, and t is the number of time units from the beginning of the scheduling (determined by the number of repetitions of a scheduling loop).

Rule TR1 is applied maximally, i.e. each cube is transformed in the form for which the rule is no longer applicable. To get rid of different but equally efficient solutions if more instances of the same machine type are free, we create only a cube where the operation is scheduled for the instance with the smallest index j. We get the following working set of cubes (variables are ordered using the ordering that turns out to be the most efficient for the ZBDD representation, as explained in Subsection 5.6):

$$X = \begin{Bmatrix} G^{3,1}_{0}B_{1,1}T^{3}_{2}R^{3,1}S^{3,1}_{1,1}O^{3}_{7,1}O^{3}_{4,2}M^{3,4,2}_{1}O^{3}_{8,3}M^{3,8,3}_{3}M_{1,1}M_{3,1}M_{1,2}FS_3; \\ G^{3,1}_{0}B_{1,1}T^{3}_{2}R^{3,1}S^{3,1}_{1,1}O^{3}_{7,1}O^{3}_{4,2}M^{3,4,2}_{1}O^{3}_{8,3}M^{3,8,3}_{3}M_{1,1}M_{3,1}FS_2; \\ G^{3,1}_{0}B_{1,1}T^{3}_{6}R^{3,1}S^{3,1}_{1,1}O^{3}_{4,1}O^{3}_{7,2}M^{3,7,2}_{1}O^{3}_{8,3}M^{3,8,3}_{3}M_{1,1}M_{3,1}M_{1,2}FS_3; \\ G^{3,1}_{0}B_{1,1}T^{3}_{6}R^{3,1}S^{3,1}_{1,1}O^{3}_{4,1}O^{3}_{7,2}M^{3,7,2}_{1}O^{3}_{8,3}M^{3,8,3}_{3}M_{1,1}M_{3,1}FS_2; \\ G^{3,1}_{0}B_{1,1}T^{3}_{6}R^{3,1}S^{3,1}_{1,1}O^{3}_{4,1}O^{3}_{7,2}M^{3,7,2}_{2}O^{3}_{8,3}M^{3,8,3}_{3}M_{1,1}M_{2,1}M_{3,1}FS_3; \\ G^{3,1}_{0}B_{2,1}T^{3}_{4}R^{3,1}S^{3,1}_{2,1}O^{3}_{7,1}O^{3}_{4,2}M^{3,4,2}_{1}O^{3}_{8,3}M^{3,8,3}_{3}M_{1,1}M_{2,1}M_{3,1}FS_3 \end{Bmatrix} \quad (6)$$

As said above, we decided to explain the scheduling algorithm for a small example with a single part P_3 to keep the number of cubes in the evolving working sets as small as possible. Therefore, only one operation (the first one in each process sequence) is scheduled in every cube of Eq. 6. In general, if the factory produces several different part types, many operations are scheduled simultaneously.

In the second scheduling phase, each cube in the working set is checked whether it includes an active part. If such a part exists, its "time to finish" is decreased by one. This procedure is realised by adapting all T-variables and can be described formally by Rule TR2.

Rule TR2: If the cube includes T_n^s and $n > 0$, then remove T_n^s and add T_{n-1}^s.

As before, the rule is applied maximally. For our example, the result of the second scheduling phase is the same as given before, only all variables T_n^3 are replaced with variables T_{n-1}^3.

In the third scheduling phase, every cube in the working set is checked whether it includes a part that has completed the operation. If such a part exists, the busy machine instance is released, and the part is moved forward into the "waiting to start the next operation" state. This procedure is expressed formally by Rule TR3:

Rule TR3: If the cube includes $R^{s,r}$, $S_{i,j}^{s,r}$, and T_0^s, then remove $R^{s,r}$, T_0^s, and $B_{i,j}$, and add $W^{s,r+1}$.

In the obtained working set of cubes, none of the parts has completed an operation. Thus, the third scheduling phase does not change anything. All three scheduling phases are then repeated, and the following working set of cubes is obtained after applying Rule TR2 for the second time:

$$X = \begin{Bmatrix} G_0^{3,1}B_{1,1}T_0^3R^{3,1}S_{1,1}^{3,1}O_{7,1}^3O_{4,2}^3M_1^{3,4,2}O_{8,3}^3M_3^{3,8,3}M_{1,1}M_{3,1}M_{1,2}FS_3\ ; \\ G_0^{3,1}B_{1,1}T_0^3R^{3,1}S_{1,1}^{3,1}O_{7,1}^3O_{4,2}^3M_1^{3,4,2}O_{8,3}^3M_3^{3,8,3}M_{1,1}M_{3,1}FS_2\ ; \\ G_0^{3,1}B_{1,1}T_4^3R^{3,1}S_{1,1}^{3,1}O_{4,1}^3O_{7,2}^3M_1^{3,7,2}O_{8,3}^3M_3^{3,8,3}M_{1,1}M_{3,1}M_{1,2}FS_3\ ; \\ G_0^{3,1}B_{1,1}T_4^3R^{3,1}S_{1,1}^{3,1}O_{4,1}^3O_{7,2}^3M_1^{3,7,2}O_{8,3}^3M_3^{3,8,3}M_{1,1}M_{3,1}FS_2\ ; \\ G_0^{3,1}B_{1,1}T_4^3R^{3,1}S_{1,1}^{3,1}O_{4,1}^3O_{7,2}^3M_2^{3,7,2}O_{8,3}^3M_3^{3,8,3}M_{1,1}M_{2,1}M_{3,1}FS_3\ ; \\ G_0^{3,1}B_{2,1}T_2^3R^{3,1}S_{2,1}^{3,1}O_{7,1}^3O_{4,2}^3M_1^{3,4,2}O_{8,3}^3M_3^{3,8,3}M_{1,1}M_{2,1}M_{3,1}FS_3 \end{Bmatrix} \tag{7}$$

Now, the last two cubes include the completed operations. Thus, Rule TR3 is applied:

$$X = \begin{Bmatrix} G_0^{3,1}B_{1,1}T_4^3R^{3,1}S_{1,1}^{3,1}O_{4,1}^3O_{7,2}^3M_1^{3,7,2}O_{8,3}^3M_3^{3,8,3}M_{1,1}M_{3,1}M_{1,2}FS_3\ ; \\ G_0^{3,1}B_{1,1}T_4^3R^{3,1}S_{1,1}^{3,1}O_{4,1}^3O_{7,2}^3M_1^{3,7,2}O_{8,3}^3M_3^{3,8,3}M_{1,1}M_{3,1}FS_2\ ; \\ G_0^{3,1}B_{1,1}T_4^3R^{3,1}S_{1,1}^{3,1}O_{4,1}^3O_{7,2}^3M_2^{3,7,2}O_{8,3}^3M_3^{3,8,3}M_{1,1}M_{2,1}M_{3,1}FS_3\ ; \\ G_0^{3,1}B_{2,1}T_2^3R^{3,1}S_{2,1}^{3,1}O_{7,1}^3O_{4,2}^3M_1^{3,4,2}O_{8,3}^3M_3^{3,8,3}M_{1,1}M_{2,1}M_{3,1}FS_3\ ; \\ G_0^{3,1}S_{1,1}^{3,1}O_{7,1}^3W^{3,2}O_{4,2}^3M_1^{3,4,2}O_{8,3}^3M_3^{3,8,3}M_{1,1}M_{3,1}M_{1,2}FS_3\ ; \\ G_0^{3,1}S_{1,1}^{3,1}O_{7,1}^3W^{3,2}O_{4,2}^3M_1^{3,4,2}O_{8,3}^3M_3^{3,8,3}M_{1,1}M_{3,1}FS_2 \end{Bmatrix} \tag{8}$$

Scheduling is continued by repeating all three phases. If a cube without M-variables and R-variables appears in the working set, it represents a solution to the scheduling problem. For the example under consideration, this happens after 14 time units, when we get the following result:

$$X = \begin{Bmatrix} G_{10}^{3,3}G_6^{3,2}G_0^{3,1}B_{3,1}T_2^3S_{1,1}^{3,1}O_{4,1}^3S_{2,1}^{3,2}O_{7,2}^3R^{3,3}S_{3,1}^{3,3}O_{8,3}^3M_{1,1}M_{2,1}M_{3,1}FS_3\ ; \\ G_{10}^{3,3}G_4^{3,2}G_0^{3,1}B_{3,1}T_2^3S_{2,1}^{3,1}O_{7,1}^3S_{1,1}^{3,2}O_{4,2}^3R^{3,3}S_{3,1}^{3,3}O_{8,3}^3M_{1,1}M_{2,1}M_{3,1}FS_3\ ; \\ G_8^{3,3}G_6^{3,2}G_0^{3,1}S_{1,1}^{3,1}O_{4,1}^3S_{1,1}^{3,2}O_{7,2}^3S_{3,1}^{3,3}O_{8,3}^3W^{3,4}M_{1,1}M_{3,1}M_{1,2}FS_3\ ; \\ G_8^{3,3}G_6^{3,2}G_0^{3,1}S_{1,1}^{3,1}O_{4,1}^3S_{1,1}^{3,2}O_{7,2}^3S_{3,1}^{3,3}O_{8,3}^3W^{3,4}M_{1,1}M_{3,1}FS_2\ ; \\ G_8^{3,3}G_2^{3,2}G_0^{3,1}S_{1,1}^{3,1}O_{7,1}^3S_{1,1}^{3,2}O_{4,2}^3S_{3,1}^{3,3}O_{8,3}^3W^{3,4}M_{1,1}M_{3,1}M_{1,2}FS_3\ ; \\ G_8^{3,3}G_2^{3,2}G_0^{3,1}S_{1,1}^{3,1}O_{7,1}^3S_{1,1}^{3,2}O_{4,2}^3S_{3,1}^{3,3}O_{8,3}^3W^{3,4}M_{1,1}M_{3,1}FS_2 \end{Bmatrix} \tag{9}$$

The last four cubes are solutions. Now, the algorithm either stops and reports one solution, reports all solutions, or merely removes all solutions and continues with scheduling until the working set of cubes is not empty. The remaining W-variable in every solution is not significant, it was only added to mark that the process sequence is finished (no process sequence for part P_3 has the fourth operation).

In the observed small example, the factory is producing only one part (P_3), and there is no competition for resources. Thus, every feasible solution has one schedule. In general, there may be several different schedules for the same feasible solution. Let us observe another problem where parts P_1, P_2, and P_3 from Fig. 1 have to be produced by using machine types M_1, M_2, M_3, and M_4 from Table 1, while the factory capacity is 2. This problem has 56 feasible solutions.

After 27 time units, the scheduling algorithm generates a working set with 763 cubes, among which there are three solutions:

$$G_{24}^{3,3}G_{20}^{3,2}G_{16}^{2,3}G_{16}^{1,3}G_{9}^{1,2}G_{8}^{2,2}G_{5}^{3,1}G_{0}^{2,1}G_{0}^{1,1}S_{4,1}^{1,1}O_{3,1}^{1}S_{2,1}^{1,2}O_{1,2}^{1}S_{4,1}^{1,3}O_{4,3}^{1}S_{2,1}^{2,1}O_{6,1}^{2}S_{4,1}^{2,2}O_{3,2}^{2}S_{2,1}^{2,3}O_{2,3}^{2}S_{2,1}^{3,1}O_{7,1}^{3}S_{4,1}^{3,2}O_{4,2}^{3}S_{4,1}^{3,3}O_{9,3}^{3}M_{2,1}M_{4,1}FC_2;$$
$$G_{24}^{3,3}G_{20}^{1,3}G_{16}^{3,2}G_{16}^{2,3}G_{9}^{1,2}G_{8}^{2,2}G_{5}^{3,1}G_{0}^{2,1}G_{0}^{1,1}S_{4,1}^{1,1}O_{3,1}^{1}S_{2,1}^{1,2}O_{1,2}^{1}S_{4,1}^{1,3}O_{4,3}^{1}S_{2,1}^{2,1}O_{6,1}^{2}S_{4,1}^{2,2}O_{3,2}^{2}S_{2,1}^{2,3}O_{2,3}^{2}S_{2,1}^{3,1}O_{7,1}^{3}S_{4,1}^{3,2}O_{4,2}^{3}S_{4,1}^{3,3}O_{9,3}^{3}M_{2,1}M_{4,1}FC_2;$$
$$G_{23}^{1,3}G_{20}^{3,3}G_{16}^{3,2}G_{16}^{2,3}G_{9}^{1,2}G_{8}^{2,2}G_{5}^{3,1}G_{0}^{2,1}G_{0}^{1,1}S_{4,1}^{1,1}O_{3,1}^{1}S_{2,1}^{1,2}O_{1,2}^{1}S_{4,1}^{1,3}O_{4,3}^{1}S_{2,1}^{2,1}O_{6,1}^{2}S_{4,1}^{2,2}O_{3,2}^{2}S_{2,1}^{2,3}O_{2,3}^{2}S_{2,1}^{3,1}O_{7,1}^{3}S_{4,1}^{3,2}O_{4,2}^{3}S_{4,1}^{3,3}O_{9,3}^{3}M_{2,1}M_{4,1}FC_2$$

All three reported schedules are based on the same feasible solution, because they have the same O-variables, S-variables, and MX-variables.

5. Implementation using zero-suppressed binary decision diagrams

Unate cube set algebra can be realised very efficiently using ZBDDs. These are one of many variants of BDDs, a relatively new computer data structure studied intensively in the 1990s [4, 7]. A computer library that implements manipulation of BDDs is called a BDD package. Only some of the available BDD packages support ZBDDs. This section gives the basic idea, describes the necessary operations, and gives algorithms used in Section 4. The implementation details of a BDD package are out of the scope of this paper. The interested readers should see [4, 5, 7], and the source code of the Biddy BDD package (available from biddy.meolic.com), which is free software used to implement all the presented algorithms [8].

5.1 Zero-suppressed binary decision diagrams

A ZBDD is a rooted binary directed acyclic graph. Every node except the leaves is called an internal node and has two descendants, called 'else' and 'then' successors, respectively. The ZBDD has an edge to the root, which is called the top edge. Each node is associated with a label. The ZBDD evaluates so that labels in internal nodes are treated as elements, and combination sets are associated with the edges. The leaves are labelled with 0 or 1, and are called terminal nodes. The label in the root is called the top label. The efficiency of ZBDDs is enhanced by fixing the order of elements along every path from the root to a leaf and minimising the graph.

The combination set can be determined precisely from a ZBDD by finding all paths starting with the top edge and leading to a terminal node 1. Each such path represents a cube. An element is included in the cube if, and only if, such a path goes through its 'then' successor. The resulting combination set is a union of the obtained cubes. Examples of some simple combination sets are given in Fig. 2.

While in the worst case, the size of ZBDDs grows exponentially with the number of elements, in many practical examples they are a very efficient data structure for the representation of combination sets. They are especially suitable for representing sets of sparse cubes, i.e., if the universal set has many elements but the cubes include only a few of them. For example, a ZBDD representing an enormous combination set including about 32×10^{12} cubes is represented with only 195 ZBDD nodes (see Section 6). The element order has a huge impact on the size of the ZBDD. To represent the same combination set with less optimal element order could require millions of ZBDD nodes.

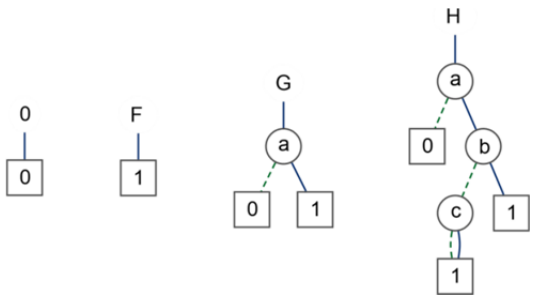


Fig. 2 ZBDDs representing different combination sets for a universal set {a,b,c}. From left to right, there are an empty set {}, a base set {{}}, set {{a}}, and set {{a},{a,b},{a,c}}

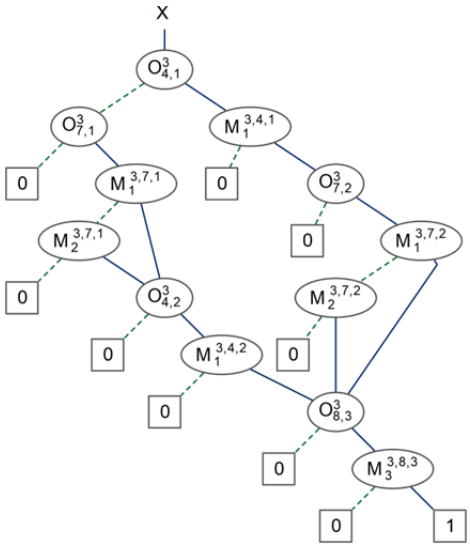


Fig. 3 A ZBDD representing a combination set (Eq. 2) with four cubes

We conclude this brief introduction of ZBDDs with Fig. 3, which gives a ZBDD representing a process plan for production of product P_3. The universal set includes five O-variables and seven M-variables. The combination set in Fig. 3 has four cubes, as described in Subsection 3.2.

5.2 Operations in unate cube set algebra

In unate cube set algebra there are all standard set operations and many special operations on cubes. We define here only those used further in the paper. In the rest of this section, a lowercase letter denotes an element, and an uppercase letter denotes a combination set. Moreover, we use the short notation of combination set introduced in Subsection 3.1.

- *Addition*, *Subtraction*, and *Intersection*: These are standard set operations. $F + G$ is the union of F and G. $F + v$ is the union of F and $\{v\}$. $F - G$ is the difference between F and G. $F \cdot G$ is the intersection between F and G. E.g., when $F = \{a; b; bc\}$, $G = \{b; c\}$, and $v = a$, $F + G = \{a; b; c; bc\}$ and $G + v = \{a; b; c\}$, $F - G = \{a; bc\}$, and $F \cdot G = \{b\}$.
- *Multiplication*: $F \times G$ is a result of all possible concatenations of two cubes in F and G. E.g., when $F = \{a; b; ac\}$ and $G = \{ab; acd\}$, $F \times G = \{a\} \times \{ab\} + \{b\} \times \{ab\} + \{ac\} \times \{ab\} + \{a\} \times \{acd\} + \{b\} \times \{acd\} + \{ac\} \times \{acd\} = \{ab\} + \{ab\} + \{abc\} + \{acd\} + \{abcd\} + \{acd\} = \{ab; abc; acd; abcd\}$.
- *Division* and *Modulo*: Quotient F/v is obtained by extraction of those cubes from F that include variable v, and removal of v from the extracted cubes. Remainder $F\%v$ is obtained by extraction of those cubes from F that do not include variable v. E.g., when $F = \{ab; bc; c\}$ and $v = b$, $F/v = \{a; c\}$ and $F\%v = \{c\}$.
- *Subset0* and *Subset1*: $Subset0(F, v)$ is exactly the same operation as *Modulo*. $Subset1(F, v)$ is obtained by extraction of those cubes from F that include variable v. E.g., when $F = \{ab; bc; c\}$ and $v = b$, $Subset0(F, v) = \{c\}$ and $Subset1(F, v) = \{ab, bc\}$.
- *Restriction*: $Restrict(F, G)$ extracts the cubes from F such that the cube is a superset of at least one cube in G. E.g., when $F = \{ab; abc; bcd; d\}$ and $G = \{abc; bc\}$, $Restrict(F, G) = \{abc; bcd\}$.
- *Change operation*: $Change(F, v)$ changes all the cubes from F such that it removes v from the cubes including it and adds v to the cubes not including it. E.g., when $F = \{ab; abc; bcd; d\}$ and $v = b$, $Change(F, v) = \{a; ac; cd; bd\}$.
- *Stretch operation*: $Stretch(F)$ extracts the cubes from F such that a proper superset of the cube is not in F. E.g., when $F = \{ab; abc; bcd; d\}$, $Stretch(F) = \{abc; bcd\}$.
- *Permitsym operation*: $Permitsym(F, n)$ extracts the cubes from F such that the cube consists of less than or equal to n elements. E.g., when $F = \{ab; abc; bcd; d\}$ and $n = 2$, $Permitsym(F, n) = \{ab; d\}$.

- *Selective multiplication*: $F \times_P^N G$ are all concatenations of two cubes in F and G such that at least one element from set P is included in a cube from G, all elements from set P are included in a cube from F if they are included in a cube from G, and also no element from N is included in a cube from F if it is included in a cube from G. E.g., when $F = \{a; b; ac\}$ and $G = \{ab; acd\}$, $F \times_{\{a\}}^{\{c\}} G = \{a\} \times_{\{a\}}^{\{c\}} \{ab\} + \{b\} \times_{\{a\}}^{\{c\}} \{ab\} + \{ac\} \times_{\{a\}}^{\{c\}} \{ab\} + \{a\} \times_{\{a\}}^{\{c\}} \{acd\} + \{b\} \times_{\{a\}}^{\{c\}} \{acd\} + \{ac\} \times_{\{a\}}^{\{c\}} \{acd\} = \{ab\} + \{\} + \{abc\} + \{acd\} + \{\} + \{\} = \{ab; abc; acd\}$.

The result of multiplication and selective multiplication with an empty set is an empty set for both cases when F or G is an empty set. Selective multiplication is not commutative.

5.3 Implementation of feasible solutions generation

To generate the set of all feasible solutions we follow mainly the approach from [1]. Algorithms are given in Fig. 4.

```
1  def createProcessPlans():
2    for s in 1..S:
3      X[s] = getSequences(s)
4      for w in 1..W:
5        for r in 1..max(s):
6          O = findO(s,w,r)
7          if O > 0:
8            z = 0
9            for i in 1..I:
10             M = findM(s,w,r,i)
11             if M > 0: z = z + M
12           z = Change(z,O)
13           X[s] = (X[s]/O)×z + (X[s]%O)
14
15 def createComprehensivePlan():
16   X = 1
17   for s in 1..S:
18     X = X × X[s]
19
20 def createMachineInstances():
21   for i in 1..I:
22     MI[i] = 0
23     for j in instances(i)..1:
24       M = findMX(i,j)
25       MI[i] = (MI[i]+1) × {M}
26
27 def createPRplan():
28   for s in 1..S:
29     for w in 1..W:
30       for r in 1..max():
31         for i in 1..I:
32           M = findM(s,w,r,i)
33           if M > 0:
34             f = Subset1(X,M)
35             g = Restrict(f,MI[i])
36             X = (f − g) × MI[i] + X%M + g
37
38 def restrictToFactoryCapacity(fc):
39   Mall = 1
40   for i in 1..I:
41     Mall = Mall × (MI[i]+1)
42   Mng = Mall − Permitsym(Mall,fc)
43   X = X − Restrict(X,Mng)
44
45 def createFeasibleSolutions():
46 # X, X[s], and MI[i] are global variables
47   createElements()
48   createProcessPlans()
49   createComprehensivePlan()
50   createMachineInstances()
51   createPRplan()
52   restrictToFactoryCapacity(FC)
```

Fig. 4 The functions involved in generating feasible solutions

In Fig. 4, S denotes the number of all part types, W denotes the number of all operation types, I denotes the number of all machine types, max(s) is the length of the longest sequence of operations to produce part P_s, and instances(i) is the number of instances of machine type M_i. We use Python-style pseudocode, where the line with a different indent than the previous one starts a new block. Function getSequences(s) returns a set of cubes such that every cube corresponds to a sequence of operations encoded with O-variables. Function createElements() creates only those elements that are included into a system. Functions findO(), findM(), and findMX() return an element with the given indices, or 0 if such an element does not exist. For specifying the set of serial numbers, we disobey the Python syntax and write $1..S$ for the set of numbers $\{1,2,\ldots,S\}$.

Function createProcessPlans() creates process plans for all parts (as Eq. 2 for P_3 in Subsection 3.2). The algorithm iterates over all values of s, w, and r. If O_w is not the r-th operation in part P_s, this combination of indices is skipped for the sake of efficiency (line 7). Othewise, a cube is added for each machine M_i that can be used to complete operation O_w. This is achieved by creating a union of adequate machines first (lines 9-11) and then calculating a product with the

subset of those cubes from the partial result z which include the particular operation (line 13). Cubes from z which do not include $O^s_{w,r}$ are kept unchanged. Please note that if operation *Subset1* were used instead of *Division*, line 12 would be removed. Because efficiency would not increase significantly, we use mathematically more elegant solution with *Division* and *Modulo* [1].

Function createComprehensivePlan() implements a simple product of process plans for all parts. This is the point where ZBDD shows its strength: while the number of cubes in the result is a product of cube numbers of the individual process plans, the size of resulting ZBDD (i.e. the number of nodes) is about a sum of sizes of involved ZBDDs (it can be even less than the sum!).

Function createMachineInstances() creates viable factory configurations for each machine as described in Subsection 3.2. For example, if machine M_1 has instances $M_{1,1}$, $M_{1,2}$, and $M_{1,3}$ then viable factory configurations are $\{M_{1,1}\}$, $\{M_{1,1}, M_{1,2}\}$, and $\{M_{1,1}, M_{1,2}, M_{1,3}\}$. Functions createPRplan() and restrictToFactoryCapacity() are the same as given and explained in [1].

5.4 Implementation of scheduling

As described in Section 4, the proposed scheduling approach is an extension of a method for generation of feasible solutions. Its implementation is described formally with the algorithms given in Fig. 5. There, functions findW(), findR(), findB(), findS(), and findT() return an element with the given indices, or 0 if such an element does not exist.

The input to the scheduling is a global variable X, which is a single ZBDD representing the set of all feasible solutions. Function schedulingInit() creates sets tr, setP, and setN (lines 9-17), which are stored in global variables and used later in schedulingPhase 1. Furthermore, it initialises scheduling procedure by adding "waiting to start the first operation" tag to all parts (lines 18-21).

Function schedulingPhase1() utilises the Selective multiplication operation for generating all possible schedules. The result of a single call to selective multiplication is a set of all cubes such that one of the parts included in the cube is scheduled to start an operation. Sometimes, more than one part can start an operation simultaneously, and, thus, there is a loop (lines 27-29). Please note that selective multiplication does not remove anything from cubes, elements are only added, and thus operation Stretch removes all schedules where Rule TR1 is not maximally applied elegantly. Searching the *full state space* where a part that can be scheduled is not required to be scheduled, can be achieved by merely omitting the Stretch operation in line 30. However, this yields a problem which is harder to calculate, while rarely bringing a better solution. The next action made in scheduling phase 1 is removing schedules where an instance with index j is occupied instead of a free instance with a lower index (lines 31-36). Finally, for cubes containing both R-variable and W-variable with the same indices (these denote processes which have just started an operation), the corresponding W-variable and M-variable are removed, and also an appropriate G-variable is added (lines 37-52). Adding G-variables is optional. By omitting line 45, the algorithm will be much more efficient, but the result of scheduling will be the set of different feasible solutions instead of a number of different schedules (if more than one instance of some machines are allowed, then the result may include several cubes for a single feasible solution).

Function schedulingPhase2() is simpler, since it only decrements the second index of every T-variable. Function schedulingPhase3() looks for cubes with T-variables that have the second index equal to zero (this denotes that an operation has been completed). For such cubes, it removes the appropriate T-variable, R-variable, and B-variable, and marks the corresponding part with the "waiting to start the next operation" tag (line 85). Function removeSolutions checks for cubes without any R-variable and M-variable. They are solutions, and the function reports and then removes them from the working set of cubes (lines 97-98). Scheduling is running until the working set of cubes is not empty (lines 103-108). Indeed, the algorithm can be modified in such a way that the scheduling is stopped after the first solution set is found.

```
def schedulingInit():
  tr = setP = setN = 0
  for s in 1..S:
    for w in 1..W:
      for r in 1..max(s):
        for i in 1..I:
          M = findM(s,w,r,i)
          if M > 0:
            for j in 1..instances(i):
              f = Change(1,M); setP = setP+{M}
              E = findW(s,r); f = Change(f,E); setP = setP+E
              E = findMX(i,j); f = Change(f,E); setP = setP+E
              E = findR(s,t); f = Change(f,E); setN = setN+E
              E = findB(i,j); f = Change(f,E); setN = setN+E
              E = findS(s,r,i,j); f = Change(f,E)
              E = findT(s,time(w,i)); f = Change(f,E)
              tr = tr + f
            f = Subset1(X,M)
            X = X − f
            f = Change(f,findW(s,r))
            X = X + f

def schedulingPhase1(step):
  Y = X ×^{setP}_{setN} tr
  X = X · Stretch(X+Y)
  f = Y
  while f != 0:
    f = f ×^{setP}_{setN} tr
    Y = Y + f
  Y = Stretch(Y)
  for i in 1..I:
    for j in 2..instances():
      f = Subset1(Y,findB(i,j))
      Y = Y − f
      f = Subset1(f,findB(i,j−1))
      X = X + f
  for s in 1..S:
    for r in 1..max(s):
      f = Subset1(Y,findR(s,r))
      Y = Y − f
      g = Subset1(f,findW(s,r))
      if g != 0:
        f = f − g
        g = Change(g,findW(s,r))
        g = Change(g,addGelement(s,r,step))
        for w in 1..W:
          for i in 1..I:
            h = Subset1(g,findM(s,w,r,i))
            g = g − h
            h = Change(h,findM(s,w,r,i))
            g = g + h
      Y = Y + f + g
  X = X + Y

def schedulingPhase2(step):
  k = 1
  complete = False
  while complete == False:
    complete = True
    for s in 1..S:
      T = findT(s,k)
      if T > 0:
        complete = False
        f = Subset1(X,T)
        X = X − f
        f = Change(f,T)
        f = Change(f,findT(s,k−1))
        X = X + f
    k += 1

def schedulingPhase3(step):
  for s in 1..S:
    for r in 1..max(s):
      g = 0
      for i in 1..I:
        for j in 1..instances(i):
          f = Subset1(X,findT(s,0))
          f = Subset1(f,findR(s,r))
          f = Subset1(f,findS(s,r,i,j))
          X = X − f
          f = Change(f,findT(s,0))
          f = Change(f,findR(s,r))
          f = Change(f,findB(i,j))
          g = g + f
      g = Change(g,findW(s,r+1))
      X = X + g

def removeSolutions(step):
  f = X
  for s in 1..S:
    for r in 1..max(s):
      f = f % findR(s,r)
      for w in 1..W:
        for i in 1..I:
          f = f % findM(s,w,r,i)
  if f != 0:
    reportSolutions(f,step)
    X = X − f

def scheduling():
  step = 1
  schedulingInit()
  while X != 0:
    schedulingPhase1(step)
    schedulingPhase2(step)
    schedulingPhase3(step)
    removeSolutions(step)
    step += 1
```

Fig. 5 Implementation of scheduling

5.5 Heuristic approach

Even if the resulting ZBDD representing a huge combination set is small, a problem may appear to create or manipulate it efficiently. Thus, a heuristic is needed to restrict the state space. Luckily, ZBDDs are very suitable for such an approach. Here, we outline two possible heuristics, the other ones, and especially the multi-objective optimization, are left for further work.

The heuristic which we can introduce the easiest is an upper bound for the *total workload* (i.e. total machine time). A weight is assigned to M-variables such that the weight of an element $M_i^{s,w,r}$ represents the time needed to finish the operation O_w on machine type M_i. All other variables have zero weights. Then, the sum of weights of all elements in a cube is the total workload for the solution corresponding to this cube. In the presented algorithms, a single ZBDD is used to represent the set of all viable solutions and, thus, a single pass over its nodes is enough to bound the total workload for all solutions simultaneously. The algorithm is given in Fig. 6. It extracts those cubes from the given set of solutions that have a total workload less than or equal to the given bound. The presented algorithm uses C-syntax, and is a typical recursive algorithm for manipulation of BDDs. Basic knowledge about BDDs is required in order to understand it. Please compare it with the algorithms given in [7] and [4]. The total workload bound is best used during a generation of feasible solutions. It should be added to the algorithm for construction of a comprehensive process plan (i.e. immediately after line 18 in Fig. 4).

```
1  ZBDD boundMachineTime(ZBDD f, unsigned int n)
2  {
3    ZBDD e,t,r;
4    unsigned int w;
5
6    if (n == 0) return emptySet;
7    if (isTerminalNode(f)) return f;
8    if (r = findInCache(f,n)) return r;
9    e = boundMachineTime(getElseSuccessor(f),n);
10   if (isMvariable(getTopLabel(f)) {
11     w = getWeight(getTopLabel(f));
12     if (n < w) n = 0; else n = n − w;
13   }
14   t = boundMachineTime(getThenSuccessor(f),n);
15   r = foaNode(getTopLabel(f),e,t);
16   storeInCache(f,n,r);
17   return r;
18 }
```

Fig. 6 A ZBDD-based algorithm for bounding total workload

Another interesting heuristic is a *makespan limit*. It can be implemented by using the same weights assigned to M-variables as already described. A sum of weights is calculated separately for each part and each machine type. If any part or machine type requires more time than the makespan limit to complete, the corresponding cube is removed from the set of solutions. Indeed, if more than one instance of a particular machine type is allowed, then the limit for this machine type should be multiplied with the number of instances. Again, a single pass recursive algorithm on the ZBDD representing the set of solutions is an efficient way to consider the given limit for makespan. We refer the interested reader to check the available source code. In contrast to the total workload bound, applying the makespan limit is not reasonable during the generation of feasible solutions. It is the most beneficial if this operation is taken before every scheduling step (i.e. immediately after line 103 in Fig. 5).

5.6 The element order in the ZBDD

The elements in ZBDD are ordered, and the order has a critical impact on its size. Therefore, we report here the order used in our implementation, which has been determined experimentally (the element in a node is considered smaller than the elements in its successors):

$B_{1,j} < B_{2,j} < ... < T_n^1 < W^{1,1} < R^{1,1} < S_{i,j}^{1,1} < O_{w,1}^1$ and $M_i^{1,w,1}$ (interleaved) $< W^{1,2} < R^{1,2} < S_{i,j}^{1,2} < O_{w,2}^1$ and $M_i^{1,w,2}$ (interleaved) $< ... < T_n^2 < W^{2,1} < R^{2,1} < S_{i,j}^{2,1} < O_{w,1}^2$ and $M_i^{2,w,1}$ (interleaved) $< ... < MX_{i,1} < MX_{i,2} < ... < FS_n < 1$

6. Results and discussion

First, the generation of all feasible solutions is observed. In Table 2 and Table 3, we report some details for the factory in Fig. 1 that produces eight different part types and has eight machine types specified. The production system involves 15 types of operations. The comprehensive process plan for this system contains 32,878,483,200,000 cubes, and is represented by our implementation with only 195 ZBDD nodes. These results can be compared with results from [1]. There, for the factory with capacity 8 and no restrictions to total workload, the authors report exactly the same number of cubes in a comprehensive process plan and the process and resource plan, but they needed 280 ZBDD nodes and 19,443 ZBDD nodes (our result is 13,072 ZBDD nodes) to represent them, respectively. Without restrictions, all the results are computed in less than a second, and, with a restriction on the maximal total workload, in just over two seconds. Thus, the exact computation times are irrelevant and not shown. When bounding the total workload to 100, the comprehensive process plan contains 245,837,448 cubes and is represented by our implementation with 1,744 ZBDD nodes. All the results have been obtained on a 3.40 GHz Intel Core i7-4770 processor with 32 GB of RAM memory.

Table 2 Size of a set of cubes representing the process and resource plans for the factory from Fig. 1

	Process and resource plan (without restrictions to total workload)		Process and resource plan (max. total workload = 100)	
Factory capacity	Number of cubes	ZBDD nodes	Number of cubes	ZBDD nodes
3	169,984	274	0	1
4	284,701,184	3,000	18488	695
5	41,207,077,120	8,193	1,526,572	5,965
6	1,365,249,188,224	11,729	26,001,900	16,898
7	14,411,349,910,656	12,925	180,702,952	24,620
8	65,501,043,610,240	13,072	642,479,776	26,190
9	164,241,617,343,104	13,041	1,398,613,308	26,181
10	272,777,626,896,512	13,011	2,158,556,924	26,162
11	343,383,824,875,136	12,994	2,620,121,648	26,145
12	364,877,105,061,888	12,985	2,747,814,784	26,136

With the help of the heuristic, we were able to count all the solutions for all the systems. The optimal lower bounds for total workload and makespan were determined experimentally by repeating the calculation with different parameters. Table 3 reports the number of all feasible solutions with the given maximal total workload, of feasible solutions yielding to the adequate schedules, and of adequate schedules for different combinations of parameters, respectively. For this system, there exist 65,501,043,610,240 feasible solutions if no restrictions on total workload and factory size are used. For scheduling, the algorithm without examining the full state space was used (see Subsection 5.4). If full state space is examined, the calculation takes much longer, and we are not able to complete it in most cases. However, in the case when the makespan is fixed to 15, and the total workload is fixed to 95, examining the full state space is feasible, and it brings an interesting result: two feasible solutions out of 9,262,892 possible ones yield to 704 adequate schedules, none of which can be found without examining the full state space.

Table 3 Results for the factory with eight parts and eight machine types (factory size is required to be exactly eight machine instances; the algorithm without examining full state space is used)

	Total workload					
	87	88	90	91	95	96
	Number of feasible solutions restricted to the total workload					
	24	672	28,112	116,376	9,262,892	22,474,516
Makespan	Number of feasible solutions yielding adequate schedules/Number of adequate schedules					
15	0/0	0/0	0/0	0/0	0/0	2/36
16	0/0	0/0	0/0	1/4	236/3,120	451/5,151
17	0/0	0/0	4/78	30/422	3,350/44,181	6,525/91,419
18	0/0	7/158	180/5,047	674/14,965	27,658/412,204	52,895/721,713
24	4/54	104/4,074	3,220/135,960	12,665/ 517,086	out of memory	out of memory

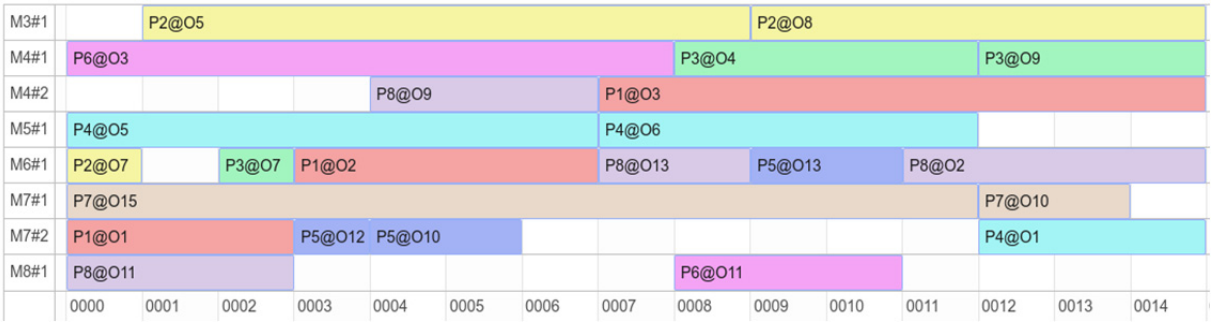


Fig. 7 Solution with makespan 15 and total workload 95

Determining the optimal solution depends on the selected goals, i.e. it is better to have a schedule with a shorter makespan, with a smaller total workload, or we want to minimise a function combining these and other objectives as well. Fig. 7 gives a Gantt chart for a solution with makespan 15 and total workload 95, which is quite hard to find and is optimal in many aspects.

To evaluate the potential of the presented method, we implemented benchmarks from [9] and [10] which have been used recently to compare different genetic and swarm algorithms [11, 12]. Without examining the full state space, we have been able to count adequate feasible solutions and schedules for all systems (see Table 4). We confirm that the most interesting solutions to these problems have been already found and reported. Still, we can present some original schedules. For example, the problem 4×5 can be solved equally well with only four machines (Fig. 8), and the problem 10×7 can be solved with makespan equal to 11 and total workload 61 (Fig. 9).

Table 4 Number of feasible solutions and schedules for benchmarks from [9] and [10] found without the full state space examination; to optimise the workload of the critical machine, maximal possible factory size was required

Problem	Makespan	Total workload	Number of feasible solutions	Number of schedules	Time to find all schedules (s)
4×5	11	32	4	4	0.36
8×8	14	77	3	4	3.14
	16	73	1	9	2.17
10×7	11	61	138	19,176	77.34
	11	62	441	44,480	205.38
	12	60	46	26,576	74.71
10×10	7	42	692	81,953	152.87
	8	41	704	818,568	630,35
15×10	11	91	60	907,570	9,003,50

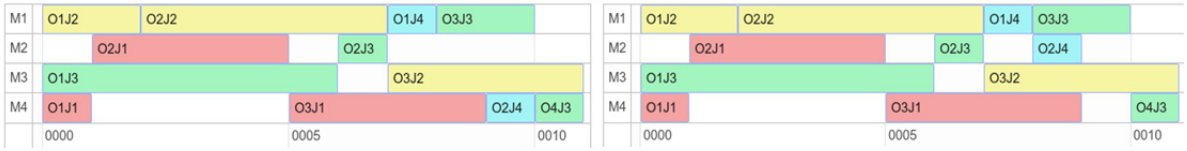


Fig. 8 There exist 4 schedules (18 in the case of the full state space exploration) based on two different feasible solutions to solve problem 4×5 with only 4 machines such that makespan is 11 and the total workload is 32

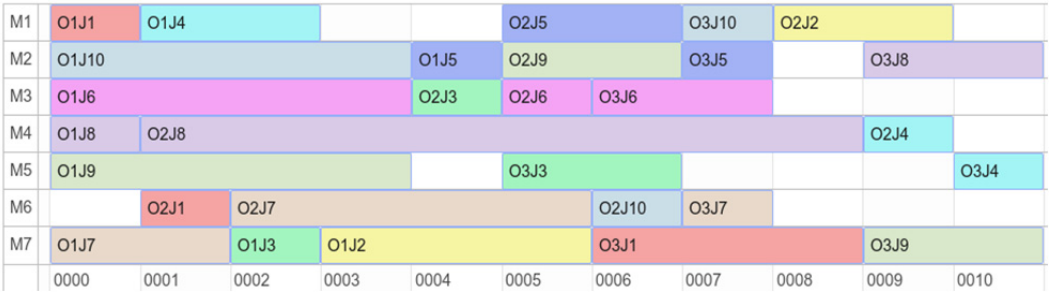


Fig. 9 Solution for problem 10×7 with makespan 11 and total workload 61

7. Conclusion

This paper proposes a novel method for generating and counting solutions of a flexible job shop scheduling problem. Feasible solutions are represented as cubes of a combination set, and the algorithm is implemented as a sequence of operations in unate cube set algebra. The efficiency of the approach is achieved by representing combination sets with ZBDDs, a data structure used typically in various methods for formal verification of systems.

In the proposed algorithm, process planning and scheduling are two sequential steps like in a typical non-linear process planning. However, in contrast to all other existing methods, the huge set of *all* feasible solutions is created in the first step. Although this looks like an unreasonable and awkward solution, ZBDDs somehow help to complete even the relatively large problems successfully. Moreover, in comparison to genetic algorithms, particle swarm optimization algorithms, and other evolutionary approaches, the obtained results are very comprehensive. For example, we have found out that the result reported in [13] is simply wrong, because for benchmark 8×8 no solution with makespan equal to 15 and total workload equal to 73 exists!

In further work, the presented algorithms can be adapted to handle additional constraints, such as the production of parts in a required order, production of several instances of each part type, and the last acceptable part delivery time. Moreover, dynamic rescheduling can be introduced to take care of eventual machine breakdowns or urgent new orders.

Acknowledgement

This work was supported in part by the Slovenian Research Agency through the Research Program Advanced Methods of Interaction in Telecommunication under Grant P2-0069.

References

[1] Takahashi, K., Onosato, M., Tanaka, F. (2014). Comprehensive representation of feasible combinations of alternatives for dynamic production planning using zero-suppressed binary decision diagram, *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, Vol. 8, No. 4, JAMDSM0061, doi: 10.1299/jamdsm.2014jamdsm0061.

[2] Takahashi, K., Onosato, M., Tanaka, F. (2015). A solution method for comprehensive solution candidates in dynamic production planning by zero-suppressed binary decision diagrams, *Transactions of the Institute of Systems, Control and Information Engineers*, Vol. 28, No. 3, 107-115, doi: 10.5687/iscie.28.107.

[3] Phanden, R.K., Jain, A., Verma, R. (2011). Review on integration of process planning and scheduling, In: Katalinic, B., (ed.), *DAAAM International Scientific Book 2011*, DAAAM International, Vienna, Austria, 593-618, doi: 10.2507/daaam.scibook.2011.49.

[4] Minato, S.-I. (1993). Zero-suppressed BDDs for set manipulation in combinatorial problems, In: *Proceedings of 30th ACM/IEEE Design Automation Conference*, Dallas, Texas, USA, 272-277.

[5] Minato, S.-I. (2001). Zero-suppressed BDDs and their applications, *International Journal on Software Tools for Technology Transfer*, Vol. 3, No. 2, 156-170.

[6] Jensen, R.A., Lauritzen, B.L., Laursen, O. (2004). Optimal task graph scheduling with binary decision diagrams, from *https://pdfs.semanticscholar.org/0fe3/36b2e5e77df4ececfa749d28752694976636.pdf?_ga=2.231505503.177204796.1540581701-1029763941.1533058815*, accessed May 5 2018.

[7] Brace, K.S., Rudell, R.L., Bryant, R.E. (1990). Efficient implementation of a BDD package. In *Proceedings of the 27th ACM/IEEE Design Automation Conference (DAC '90)*, New York, USA, 40-45. doi: 10.1145/123186.123222.

[8] Meolic, R., (2012). Biddy – A multi-platform academic BDD package, *Journal of Software*, Vol. 7, No. 6, 1358-1366, doi: 10.4304/jsw.7.6.1358-1366.

[9] Kacem, I., Hammadi, S., Borne, P. (2002). Pareto-optimality approach for flexible job-shop scheduling problems: Hybridization of evolutionary algorithms and fuzzy logic, *Mathematics and Computers in Simulation,* Vol. 60, No. 3-5, 245-276, doi: 10.1016/S0378-4754(02)00019-8.

[10] Kacem, I., Hammadi, S., Borne, P. (2002). Approach by localization and multiobjective evolutionary optimization for flexible job-shop scheduling problems*, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* Vol. 32, No. 1, 1-13, doi: 10.1109/TSMCC.2002.1009117.

[11] Chaudry, I.A., Khan, A.M., Khan, A.A. (2013). A genetic algorithm for flexible job shop scheduling, In: *Proceedings of the World Congress on Engineering Vol. I*, London, U.K., 703-708.

[12] Yu, M.R., Yang, B., Chen, Y. (2018). Dynamic integration of process planning and scheduling using a discrete particle swarm optimization algorithm, *Advances in Production Engineering & Management*, Vol. 13, No. 3, 279-296, doi: 10.14743/apem2018.3.290.

[13] Zhang, H., Gen, M. (2005). Multistage-based genetic algorithm for flexible job-shop scheduling problem, *Journal of Complexity International*, Vol. 11, 223-232.

APEM journal
Advances in Production Engineering & Management
Volume 13 | Number 4 | December 2018 | pp 389–404
https://doi.org/10.14743/apem2018.4.298
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Functional objectives decision-making of discrete manufacturing system based on integrated ant colony optimization and particle swarm optimization approach

Xu, W.[a], **Yin, Y.**[b,*]

[a]School of Management, Shenyang University of Technology, Shenyang, P.R. China
[b]ANU College of Business and Economics, The Australian National University, Australia

ABSTRACT

In order to obtain a decision model with universality, the manufacturing unit was regarded as the most basic carrier for the functional objectives of the manufacturing system. This paper has established the functional objective decision model of discrete manufacturing system by characterizing the manufacturing objectives of cost, efficiency, quality, time, agility and greenness, and has introduced the concept of coordination degree between manufacturing units. In weight calculation, the model could balance the importance of the functional objectives required by the customer and the producer. Moreover, according to the NP-hard characteristics of the model, ant colony algorithm and particle swarm optimization (ACO-PSO) algorithm was designed to solve the problem. The feasibility and validity of the algorithm were verified by simulation examples, which could promise the experimental results more satisfactory than the traditional genetic algorithm. In addition, the model can provide more choices for decision-making of functional objectives in discrete manufacturing systems by adjusting the fitness value.

ARTICLE INFO

Keywords:
Discrete manufacturing;
Functional objectives;
Decision-making;
Ant colony optimization (ACO);
Particle swarm optimization (PSO)

**Corresponding author:*
lu6400269@anu.edu.au
(Yin, Y.)

Article history:
Received 25 August 2018
Revised 11 November 2018
Accepted 30 November 2018

1. Introduction

Manufacturing system is the core subsystem of enterprises and the fundamental driving force for sustainable development of enterprises [1]. Its competitiveness directly determines the competitiveness of product market, so enterprises have to improve the manufacturing system to meet the fluctuation of market and customer demand. The companies can ensure their competitive advantages and their continued effectiveness in a complex market environment through this method [2]. In previous investigations, the manufacturing, in which only 5 % of the total manhour cost was invested, was used in enterprises, but the total cost of the product was increased by 12 times [3-4]. Therefore, good adaptability and competitiveness of enterprises cannot do without flexible adjustment of the manufacturing system [5]. Besides, practice has proved that, the main way to improve the flexible adaptability and competitiveness of enterprises is to build a manufacturing system, and it could meet market demands [6-7]. Furthermore, in order to ensure the realization of environmental needs, enterprises should constantly adjust their manufacturing systems with the competition environment changing [8].

Objective decision-making function of the manufacturing system is to obtain the ideal products for users through a good manufacturing system, fundamentally provide strategic competitiveness, leverage manufacturing system functions, meet multiple technical and economic indi-

cators, and identify the best comprehensive solution among series of scenarios [9]. And at the core stage of forming enterprise competitiveness, the objective of manufacturing system function is the ultimate source of ensuring product and service quality, because its advantages and disadvantages directly determine the intrinsic quality of products and the external level of service, which mainly manifested in manufacturing, use and maintenance process, simultaneously, its advantages and disadvantages affect the production and service capabilities of products as well as follow-up behaviours [10-11]. In recent years, industrial technology has flourished, and users had higher requirements for products and equipment, which include quality, function, structure, etc. [12]. So, enterprises have been inseparable from the objective of manufacturing system functions during operation [13].

2. Problem descriptions and hypothesis

2.1 Problem descriptions

It can be concluded that the main functional objectives of discrete manufacturing systems include cost, efficiency, quality, time, agility and green attributes, furthermore, in order to meet customer demand and corporate competitive strategies, these may also be expanded in real-world builds [6]. Therefore, the decision-making factors are complicated. And on the other hand, in the case of limited enterprise resources, various functional objectives may be modified according to certain principles to meet the focus shifts in the production process, and there may even be serious conflicts between functional objectives. Therefore, enterprises must make scientific and effective decisions. When companies pursue the following objectives that include high efficiency, low cost, good quality, prompt delivery, high agility, greenness and coordination among the above objectives, for any of the manufacturing subtasks in the manufacturing unit of Fig. 1, each candidate unit is equivalent to a selection in *D*. That is, the manufacturing unit has to select a certain direction in each two-way arc in *D*. and if it is directed acyclic, it corresponds to an optimal service configuration that belongs to the *k*-th manufacturing subtask; Similarly, when the service execution nodes are all determined which correspond to the subtasks in the manufacturing task chain, a complete directed acyclic graph called Hamilton loop can be obtained.

There are four structures in the model which include series, parallel, selection and cycle, as shown in Fig. 2.

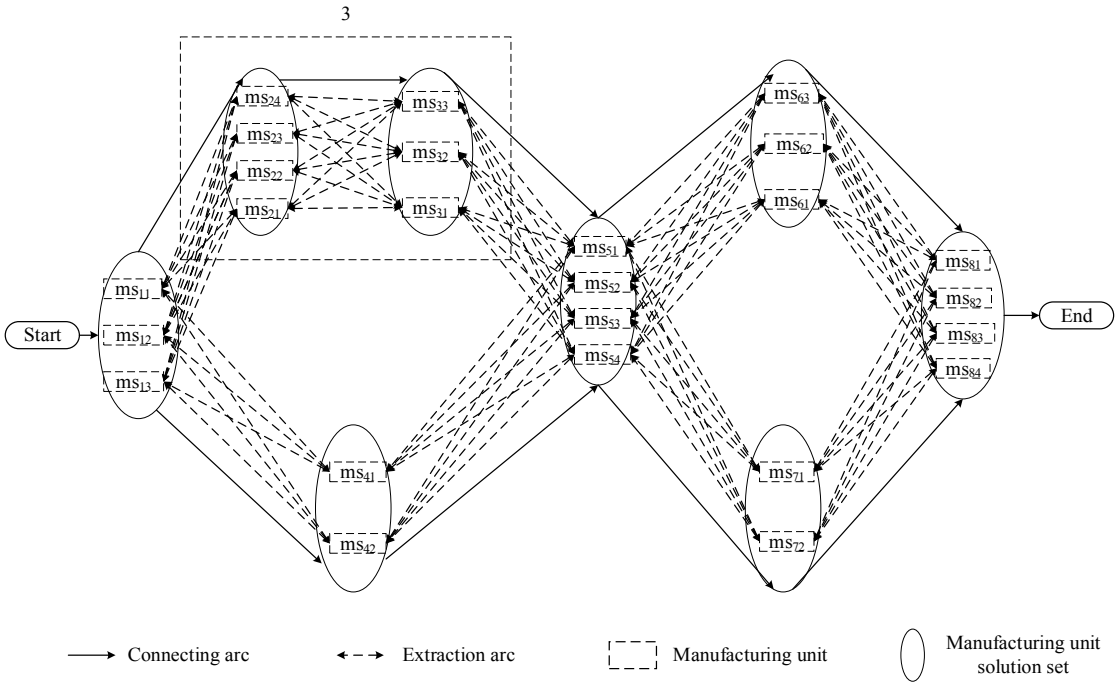
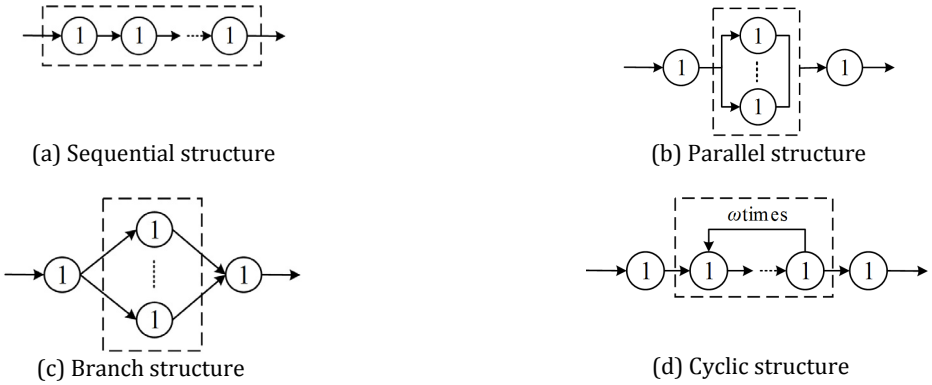


Fig. 1 Disjunctive graph of function decision making for discrete manufacturing system



(a) Sequential structure (b) Parallel structure

(c) Branch structure (d) Cyclic structure

Fig. 2 Four kinds of subtask combination structures

There is a cyclic structure in sub-task 2 and 3 which cannot meet Hamilton loop requirements, while in order to meet the requirements, three cycles need to be converted into three series. In fact, both branch selection structure and parallel structure can be converted into series structure. And it was expressed as $\{st_1 \rightarrow st_2/st_3 \rightarrow st_4 \rightarrow 3[st_5 \rightarrow st_6] \cup st_7 \rightarrow st_8\}$.

2.2 Problem hypothesis

The manufacturing unit could optimize the functional objective direction and strength of the decomposable task in a certain predictable time, and which is called discrete manufacturing system function objective decision. However, the manufacturing unit may also be self-interested, exclusive, and competitive in the actual operation process, which may cause damage to the organic whole of the functional objectives of the manufacturing system. So in order to construct a functional objective decision model with certain universality, we give such simplifications and assumptions:

- Each sub-task in the manufacturing task chain could be realized by the same type of manufacturing unit, but the premise is that it has the same processing capability, processing technology and processing method. However, each manufacturing unit has its own characteristics, which are mainly reflected in the functional objectives not in manufacturing process. Besides, a sub task could only be finished by one of the best manufacturing units.
- The manufacturing unit has differences in time, quality, cost, agility, efficiency, and green functionality, each of which can be independently measured within the optimal range of objectives. Moreover, we can comprehensively measure the functional objectives of manufacturing units by using the relevant duplicate weights, which called comprehensive evaluation, and it could provide quantitative support for the final multi-objective decision-making.
- The calculation of transportation costs and transfer time between manufacturing units is defined by a coordination factor, and it could be obtained by experts or engineering designers with reference to transportation costs and transfer times.
- There are four kinds of interconnection among manufacturing units: serial, parallel, selective and cyclic, and all of them can be expressed in series.
- The manufacturing unit function has certain robustness, and which indicates that the manufacturing system has sufficient resources and capabilities within a certain range, and it can avoid time conflicts for multiple manufacturing subtasks in the use of manufacturing units.

3. Model establishment

3.1 Objective functions

Cost objective function

For discrete manufacturing enterprises, manufacturing is the creative process of enterprise value, and the cost of manufacturing directly affects the profit and core competitiveness of discrete manufacturing enterprises. Therefore, the cost factor most concerned is manufacturing cost generated by production activities in workshops. Thus the total cost objective function of discrete manufacturing system is defined as:

$$E = \sum_{i=1}^{a}\sum_{j=1}^{k_j} H_{ij}e_{ij} + \sum_{l=1}^{b} max\left(\sum_{j=1}^{k_j} H_{ij}e_{ij}, \sum_{j=1}^{k_j} H_{(i+1)j}c_{(i+1)j}\right) + p_i\sum_{i=1}^{c}\sum_{j=1}^{k_j} H_{ij}e_{ij} + \lambda\sum_{i=1}^{d}\sum_{j=1}^{k_j} H_{ij}e_{ij} \quad (1)$$

$$H_{ij} = \begin{cases} 1 & \text{The } i\text{- th sub task is made by the manufacturing unit } ms_j \\ 0 & \text{The } i\text{- th sub task is not made by the manufacturing unit } ms_j \end{cases}$$

In Eq. 1, c_{ij} is the manufacturing cost of the *j*-th manufacturing unit in the *i*-th subtask in the manufacturing system, a is the number of serial tasks in the manufacturing unit, b is the number of parallel manufacturing units in a manufacturing system, c is the number of manufacturing units in the manufacturing system, d is the number of circulating manufacturing units in the manufacturing system, k_p is the number of manufacturing unit subtasks in the four manufacturing categories, p_i is selection probability, λ is the number of cycles for manufacturing subtasks.

Manufacturing efficiency objective function

Manufacturing efficiency generally refers to the quantity of qualified products produced per unit of time (such as one hour, one day and night), per unit of capacity of a manufacturing unit (such as a machine tool or an automated production line) or equipment (such as per cubic meter of blast furnace volume). So the increase in manufacturing efficiency showed that the manufacturing system makes full use of resources. Thus the efficiency objective function in the function objective of discrete manufacturing system is defined as:

$$E = \frac{1}{a}\sum_{i=1}^{a}\sum_{j=1}^{k_j} H_{ij}e_{ij} + \frac{1}{b}\sum_{l=1}^{b}\left(\sum_{j=1}^{k_j} H_{ij}e_{ij}, \sum_{j=1}^{k_j} H_{(i+1)j}e_{(i+1)j}\right) + p_i\frac{1}{c}\sum_{i=1}^{c}\sum_{j=1}^{k_j} H_{ij}e_{ij} + \frac{1}{d}\lambda\sum_{i=1}^{d}\sum_{j=1}^{k_j} H_{ij}e_{ij} \quad (2)$$

$$H_{ij} = \begin{cases} 1 & \text{The } i\text{- th sub task is made by the manufacturing unit } ms_j \\ 0 & \text{The } i\text{- th sub task is not made by the manufacturing unit } ms_j \end{cases}$$

In Eq. 2, e_{ij} is the manufacturing efficiency of the *j*-th manufacturing unit in the *i*-th sub-task in the discrete manufacturing system.

Manufacturing quality objective function

Manufacturing quality index is mainly used to evaluate the production capacity of machine tools equipment and the level of machining accuracy, such as production processing accuracy, product qualification rate, surface roughness and so on. Manufacturing quality of a manufacturing unit should be the product of the qualified rate of each process within the unit, that is, the manufac-

turing quality should be optimized as far as possible. Thus the objective function of manufacturing quality of discrete manufacturing system is defined as:

$$Q = \prod_{i=1}^{a}\prod_{j=1}^{k_j} H_{ij}q_{ij} + min\prod_{i=1}^{b}\left(\prod_{j=1}^{k_j} H_{ij}q_{ij}, \prod_{j=1}^{k_j} H_{(i+1)j}q_{(i+1)j}\right) + p_i\prod_{i=1}^{c}\prod_{j=1}^{k_j} H_{ij}q_{ij} + \left(\prod_{i=1}^{d}\prod_{j=1}^{k_j} H_{ij}q_{ij}\right) \tag{3}$$

$$H_{ij} = \begin{cases} 1 & \text{The } i\text{- th sub task is made by the manufacturing unit } ms_j \\ 0 & \text{The } i\text{- th sub task is not made by the manufacturing unit } ms_j \end{cases}$$

In Eq. 3, q_{ij} is the quality pass rate of the *j*-th manufacturing unit in the *i*-th subtask of a discrete manufacturing system.

Time objective function

The technological process for each order is fixed. The raw materials have gone through several processes in the workshop before becoming end products. And the processing cycle is sum of the process time of each manufacturing, in which involves handling time, waiting time, and processing time. Thus the total running time objective function of the discrete manufacturing system functional objective is defined as:

$$T = \sum_{i=1}^{a}\sum_{j=1}^{k_j} H_{ij}t_{ij} + \sum_{l=1}^{b} max\left(\sum_{j=1}^{k_j} H_{ij}t_{ij}, \sum_{j=1}^{k_j} H_{(i+1)j}t_{(i+1)j}\right) + p_i\sum_{i=1}^{c}\sum_{j=1}^{k_j} H_{ij}t_{ij} + \lambda\sum_{i=1}^{d}\sum_{j=1}^{k_j} H_{ij}t_{ij} \tag{4}$$

$$H_{ij} = \begin{cases} 1 & \text{The } i\text{- th service node selects service } ms_j \text{ to provide services} \\ 0 & \text{The } i\text{- th service node doesn't select service } ms_j \text{ to provide services} \end{cases}$$

In Eq. 4, t_{ij} is the execution time of the *j*-th manufacturing unit of the *i*-th subtask in the discrete manufacturing system.

Manufacturing agility objective function

The manufacturing agility of discrete manufacturing systems is primarily measured by the average level of agility of each manufacturing unit, which is to make the overall agility of the manufacturing system optimal as far as possible. So the objective function of manufacturing agility of discrete manufacturing system is defined as:

$$A = \left(\prod_{i=1}^{a}\prod_{j=1}^{k_j} H_{ij}a_{ij}{}'\right)^{\frac{1}{a}} + \sum_{l=1}^{b} min\left(\prod_{j=1}^{k_i} H_{ij}q_{ij}, \prod_{j=1}^{k_j} H_{(i+1)j}a'_{(i+1)j}\right)^{\frac{1}{b}} + p_i\left(\prod_{i=1}^{c}\prod_{j=1}^{k_j} H_{ij}a'_{ij}\right)^{\frac{1}{c}} + \left(\prod_{1}^{\lambda}\prod_{i=1}^{d}\prod_{j=1}^{k_j} H_{ij}a'_{ij}\right)^{\frac{1}{d}} \tag{6}$$

$$H_{ij} = \begin{cases} 1 & \text{The } i\text{- th sub task is made by the manufacturing unit } ms_j \\ 0 & \text{The } i\text{- th sub task is not made by the manufacturing unit } ms_j \end{cases}$$

In Eq. 5, a'_{ij} is the agility of the *j*-th manufacturing unit in the *i*-th subtask in a discrete manufacturing system.

Green function

The greenness of the manufacturing unit refers to the evaluation of environmental indicators in the manufacturing process under current environmental requirements, such as carbon emissions and material loss. So the objective function of manufacturing agility of discrete manufacturing system is defined as:

$$T = \frac{1}{a}\sum_{i=1}^{a}\sum_{j=1}^{k_j} H_{ij}g_{ij} + \sum_{l=1}^{b} max\left(\sum_{j=1}^{k_j} H_{ij}t_{ij}\,,\sum_{j=1}^{k_j} H_{(i+1)j}e'_{(i+1)j}\right) + p_i\frac{1}{c}\sum_{i=1}^{c}\sum_{j=1}^{k_j} H_{ij}g_{ij} + \frac{1}{d}\prod_{1}^{\lambda}\prod_{i=1}^{d}\prod_{j=1}^{k_j} H_{ij}g_{ij} \tag{6}$$

$$H_{ij} = \begin{cases} 1 & \text{The } i\text{- th sub task is made by the manufacturing unit } ms_j \\ 0 & \text{The } i\text{- th sub task is not made by the manufacturing unit } ms_j \end{cases}$$

In Eq. 6, e'_{ij} is the greenness of the j-th manufacturing unit in the i-th subtask in the discrete manufacturing system.

Manufacturing unit function objective total objective function

In actual production and operation, enterprise pursued such objectives which include low cost, high efficiency, perfect quality, short time, agility and strong green. However, these objectives are often interrelated and conflict with each other. When a manufacturing unit was selected, each sub-objective has its own optimization criteria, and they could also restrict each other, so it is almost impossible to find a group of manufacturing units that can satisfy these requirements simultaneously, which includes Min T, Max Q, Min C, Max F, Max E, Max G. Thus the corresponding weights should be assigned according to the relative importance of each sub objective.

$$minZ = w_1\,T + w_2\,C + w_3\,(1-Q) + w_4\,(1-A) + w_5\,(1-E) + w_6(1-G) \tag{7}$$

Coordination function between manufacturing units

$$R = \sum_{i=1}^{n-1}\sum_{j=1}^{k_j}\sum_{q=1}^{k_{j+1}} H_{ij}\,H_{(i+1)g}r_{ij,(i+1)q} \tag{8}$$

$$H_{ij} = \begin{cases} 1 & \text{The } i\text{- th sub task is made by the manufacturing unit } ms_j \\ 0 & \text{The } i\text{- th sub task is not made by the manufacturing unit } ms_j \end{cases}$$

In Eq. 8, $r_{ij,(i+1)q}$ is the coordination relationship between the j-th manufacturing unit in the i-th subtask in the discrete manufacturing system and the q-th manufacturing unit in the (i+1)-th subtask, and the value range is [0,1], and the greater the figure, the higher the coordination.

Functional objective decision model of discrete manufacturing system

Discrete manufacturing system function objective decision model is equivalent to the sum of the coordination values between the manufacturing unit functional objective and the manufacturing unit, as shown in Eq. 9.

$$minZ' = w_1\,T + w_2\,C + w_3\,(1-Q) + w_4\,(1-A) + w_5\,(1-E) + w_6(1-G) + w_7(1-R) \tag{9}$$

In Eq. 9, W is used as a weight vector, determined by both customers and producers.

The importance of cost (C), efficiency (E), quality (Q), time (T), agility (A), greenness (G) and coordination (R) has different dimensions and scales, so it is impossible to compare these indicators directly. Therefore, it is necessary to normalize them to eliminate the dimensional differences and obtain comparable scales. And T and C are the cost indicators z, and Q, A, E and R are

benefit indicators y. Besides, different indicators are treated differently in the normalization process, as follows:

$$Z_{ij} = \frac{V_{ij} - \min\limits_{1\le j\le k}\{V_{ij}\}}{\max\limits_{1\le j\le k}\{V_{ij}\} - \min\limits_{1\le j\le k}\{V_{ij}\}} \tag{10}$$

$$Z_{ij} = \frac{\max\limits_{1\le j\le k}\{V_{ij}\} - V_{ij}}{\max\limits_{1\le j\le k}\{V_{ij}\} - \min\limits_{1\le j\le k}\{V_{ij}\}} \tag{11}$$

After normalizing the time (T), quality (Q), cost, (C), agility (A), efficiency (E), green (G) and coordinated (R), the general objective function of discrete manufacturing system is defined as follows in Eq. 12:

$$\begin{aligned} minZ'' = w_1\, T + w_2\, C' + w_3\,(1 - Q') + w_4\,(1 - A') + w_5\,(1 - E') + w_6(1 - G') \\ + w_7(1 - R') \end{aligned} \tag{12}$$

$$w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 = 1$$

Weights $w_1, w_2, w_3, w_4, w_5, w_6,$ and w_7 are the importance of cost (C), efficiency (E), quality (Q), time (T), agility (A), greenness (G), and coordination (R), respectively.

3.2 Constraints and decision variables

The decision-making process is that each manufacturing unit provides a functional objective carrier and integrates the discrete manufacturing system, which has the most coordination of manufacturing units. Therefore, the decision has the following corresponding constraints:

Manufacturing time constraint

The total manufacturing time of the manufacturing system cannot be greater than the latest delivery time, which is:

$$T_{max} \geq T$$

T_{max} is the latest delivery time in the manufacturing system.

Manufacturing quality constraint

The manufacturing quality of the manufacturing system cannot be less than the minimum quality requirement:

$$Q_{max} \leq Q$$

Q_{max} is the minimum manufacturing quality acceptable to the manufacturing system.

Manufacturing cost constraint

The manufacturing quality of the manufacturing system cannot be higher than the maximum cost requirement.

$$C_{max} \geq C$$

C_{max} is the highest cost that a manufacturing system can afford.

Manufacturing agility constraint

The agility of manufacturing system must be higher than the minimum system agile requirement, as follows:

$$A_{min} \leq A$$

R_{min} is the minimum system agility acceptable to a manufacturing system

Manufacturing efficiency constraint

The manufacturing efficiency of the manufacturing system cannot be less than the minimum efficiency, as follows:

$$E_{min} \leq E$$

E_{min} is the minimum efficiency acceptable to the manufacturing system.

Green constraint

The greenness of the manufacturing system cannot be less than the minimum greenness, as follows:

$$G_{min} \leq G$$

G_{min} is the minimum greenness requirements acceptable to the manufacturing system

Decision variable constraint

The decision variable constraint is represented by a manufacturing sub task, and which can only be completed by one type of manufacturing unit, one has the following expression:

$$\forall_i \in \{1,2,\cdots,n\}, j \in \{1,2,\cdots,g_j\}$$

$$H_{ij} = \begin{cases} 1 & \text{The } i\text{- th sub task is made by the manufacturing unit } ms_j \\ 0 & \text{The } i\text{- th sub task is not made by the manufacturing unit } ms_j \end{cases}$$

4. Solution of model

4.1 Part of ACO

We set the number of manufacturing units to *N*, the total number of ants to *N*, and the degree of coordination between unit *i* and unit *j* is $R_{ij}(i \neq j)$. The objective function can be described that the sum of the functional objectives of each unit and the degree of coordination between different units, such as Eq. 9. The bigger the value, the better. Where the manufacturing system task sequence is expressed as {st_1, st_2, ..., st_n}, which gives a full array of all tasks.

The basic idea of ant colony algorithm: ants find the path randomly at first time, and release the same amount of pheromone on the path. We define $\tau_{ij}(t)$ to represent the residual pheromone content on the line unit *i* and unit *j* at time *t*. The pheromone content of each path at the initial moment is $\tau_{ij}(0) = \tau_0$. During the crawling process, ant $ant_k(k = 1, 2, \ldots, N)$ uses the Tabu Table $tabu_k$ to record the currently traversed unit, and uses the set $allowed_k$ to record the next alternative unit, i.e. $allowed_k = \{C - tabu_k\}$.

$p_{ij}^k(t)$ represents the state transition probability that ant ant_k is transferred from unit *i* to unit *j* at time *t*, where η_{ij} indicates the heuristic information from unit *i* to unit *j* (generally, $\eta_{ij} = 1/R_{ij}$. The parameter α represents the relative importance of the amount of residual information on the path from *i* to *j*, and the parameter β is the relative importance of the heuristic information, as shown by the Eq. 13.

$$p_{ij}^k(t) = \begin{cases} \dfrac{[\tau_{ij}(t)]^{\alpha}[\eta_{ij}]^{\beta}}{\sum_{s \in allowed_k}[\tau_{is}(t)]^{\alpha}[\eta_{is}]^{\beta}}, & \text{if } j \in allowed_k \\ 0, & j \in tabu_k \end{cases} \tag{13}$$

Eq. 14 represents the pheromone volatilization rate over time *N*; $\Delta\tau_{ij}$ represents the sum of the pheromone increments on path from *i* to *j* in this iteration, as shown in Eq. 15; Q represents the total pheromone content released by the ant cycle; L_k indicates the length of the path taken by ant_k in this tour; $\Delta\tau_{ij}^k$ represents the pheromone content left by ant_k on path in this iteration, as shown in Eq. 16.

$$\tau_{ij}(t+n) = \rho \cdot \tau_{ij}(t) + \Delta\tau_{ij} \tag{14}$$

$$\Delta\tau_{ij} \sum_{k=1}^{m} \Delta\tau_{ij}^{k} \tag{15}$$

$$\Delta\tau_{ij}^{k} = \begin{cases} \frac{Q}{L}, & \text{if } ant_k \text{ pass path } i \text{ to } j \text{ from time } t \text{ to } t+1 \\ 0, & \text{others} \end{cases} \tag{16}$$

4.2 Part of PSO

Particle swarm optimization is a swarm intelligence algorithm that simulates the predation behavior of flocks. In the process of optimization, each potential best solution is related to the velocity of particle motion. According to the historical experience of the particle and the experience of neighbors, the velocity and direction are adjusted to approach the best solution.

We set the position and velocity of the *i*-th particle in the particle group are $x_i(x_{i1}, x_{i2})$ and $v_i(v_{i1}, v_{i2})$ respectively. $pbest_i(p_{i1}, p_{i2})$ *r*epresents the best position searched by the *i*-th particle. $gbest_i(p_{i1}, p_{i2})$ represents the best position of all particles searched. according to Eqs. 17 and 18, we can update the velocity and position of the particles:

$$v_{i+1} = \omega * v_i + c_1 * \zeta_1(pbest_i - x_i) + c_2 * \zeta_2(gbest_i - x_i) \tag{17}$$

$$x_{i+1} = x_i + v_{i+1} \tag{18}$$

where the learning factors c_1 and c_2, respectively, indicate that the particles have self-optimizing ability and swarm intelligence ability to approximate the individual and group optimal solutions. ζ_1 and ζ_2 are the random numbers between [0, 1]. Particle velocity needs to be limited, if $|v_i| > v_{max}$, then $|v_i| = v_{max}$. ω is the inertia weight, which indicates how much the particle continues the current velocity, and its value is reasonably selected so that the particle has balanced local and global search ability.

4.3 Algorithm flow

The ACO-PSO hybrid algorithm model overcomes the shortcomings of both Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO) algorithms, and it combines the advantages of the two algorithms. Furthermore, its basic construction ideas are as follows: The first step is to generate an initial solution set of the problem by the strong global search ability of the particle swarm algorithm and its randomness. And then according to the primary selection result, the initial pheromone matrix of the ant colony algorithm is constructed. Next, it is to find the optimal solution of the problem, which applies the strong optimization ability and positive feedback of the ant colony algorithm. Moreover, in the optimal configuration model, in order to alleviate the occurrence of ant colony and stagnation, limited ant resources should be allocated to populations, and it has higher search rates. The algorithm structure is shown in Fig. 3.

4.4 The steps of algorithm

Step 1: Set the particle swarm parameters and initialize the particle swarm, that is, randomly generate N_p particle velocities and positions.
Step 2: Initialize the pheromone and parameters of each ant; randomly place N_{sa} ants on *N* units.
Step 3: Set each ant traverse all the units. The ant selects the passing unit by the concept of (13), updates the path length of N_{sa} ants, and calculates the optimal path of each ant as the fitness value of N_p particles.
Step 4: According to the fitness value of the particle, we update each particle's individual optimal solution and group optimal solution according to Eqs. 17 and 18, and obtain the velocity and position of the particle.
Step 5: Apply the global pheromone update rule.
Step 6: If the maximum number of iterations is reached, then the process ends. Otherwise, return to step 4.

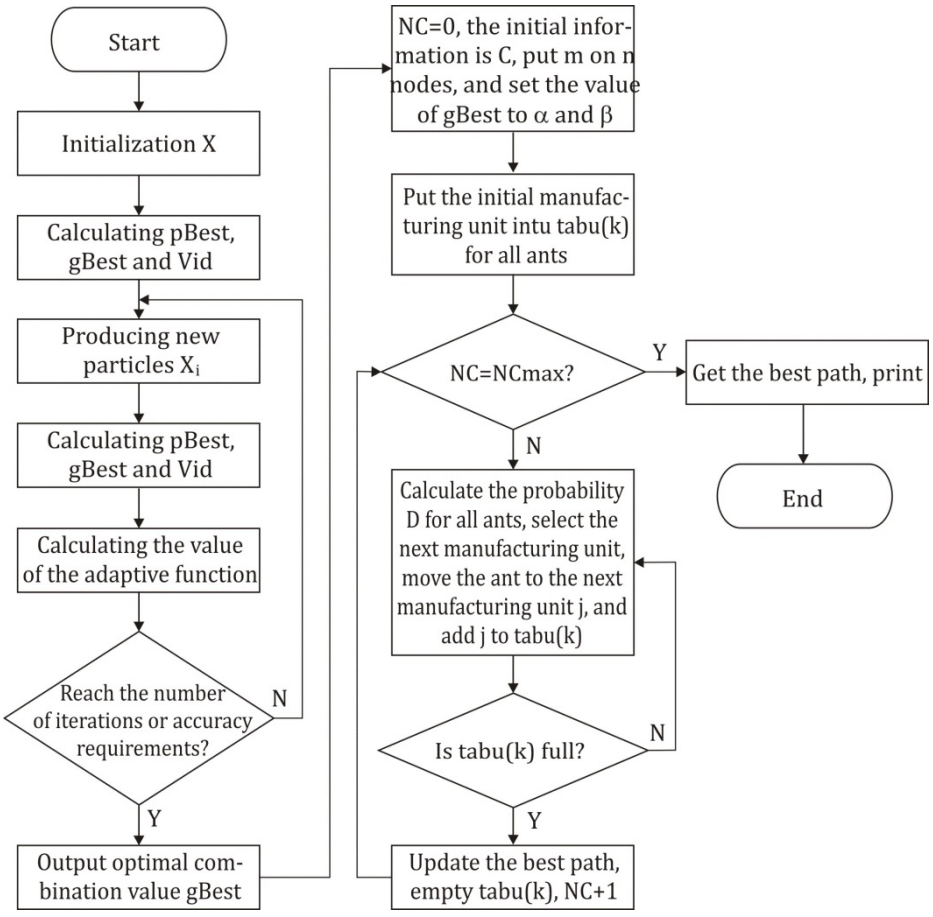


Fig. 3 Hybrid PSO-ACO algorithm flow

5. Results and discussion

5.1 The initial data

The process of functional objective decision making for discrete manufacturing systems is as follows: Starting from the upstream sub-task st_1, each sub-task st_1 passes through the downstream sub-task st_1 in the direction of directed connection arc, input and output connection arc r_{ij}. We select variable function allocate (ms_{ij}) respectively according to the rules of adjacency combination from the corresponding manufacturing units. A manufacturing unit is selected to participate in the formation of a discrete manufacturing system, which has formed satisfactory functional objectives, as shown in Fig 4. The result of discrete manufacturing system decision is that, the preferred manufacturing units has been selected from each sub-task, and then combined them to form a complete organic whole E^*, which can be formally expressed as:

$$E^* = \left\{\bar{P}\left(ms_{ij}\right),\ ms_{ij},\ \bar{N}\left(ms_{ij}\right)\right\}$$

According to Fig. 4, an optimal decision-making scheme can be established, and it is corresponded to the task chain which consists of manufacturing unit, as shown in Table 1.

Functional objective decision-making of manufacturing system is the process of screening and integrating manufacturing units for each manufacturing sub-task in the manufacturing task chain. Besides it can make full use of manufacturing units based on the specific manufacturing system construction scheme, meanwhile it could optimize the allocation of resources within enterprises, and then achieve the organic combination of various manufacturing units under different manufacturing modes which can help form the dynamic core competitiveness, and at last maximize the functional objectives of manufacturing resources utility.

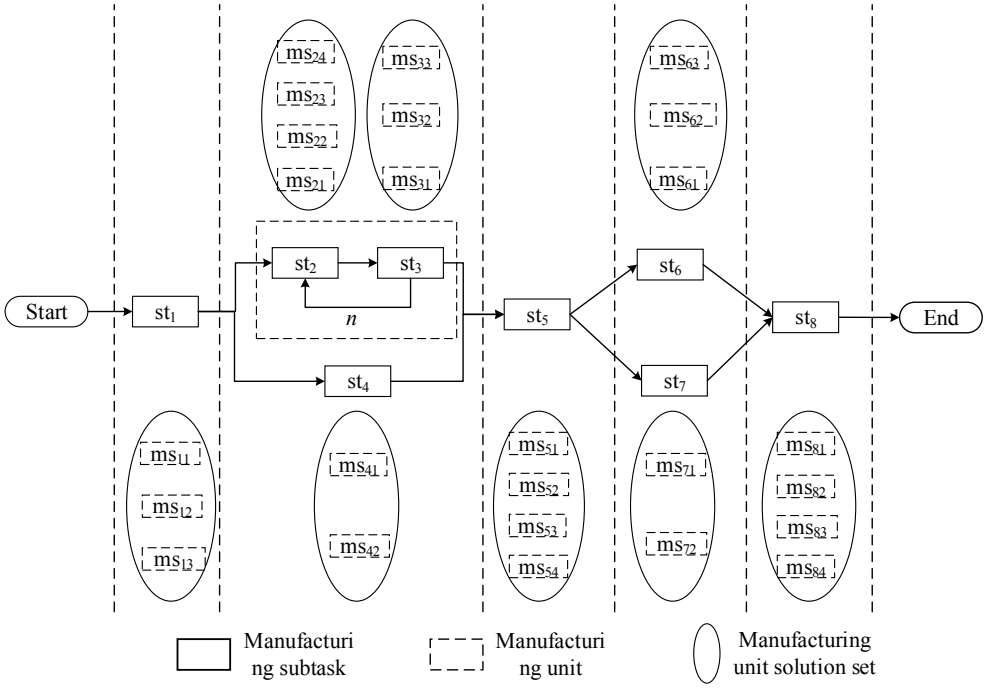


Fig. 4 Sketch map of function decision of discrete manufacturing system

Table 1 Functional components of discrete manufacturing system

Subtask	Manufacturing unit	C	E	T	Q	A	G
ts_1	ms_{11}	23.4	10	13.4	17.5	8.9	7.1
	ms_{12}	25.6	12.5	16.5	11.5	8.9	14.1
	ms_{13}	24.5	9.8	21.5	8.7	14.5	6.8
ts_2	ms_{21}	25.4	18.5	9.8	15.2	9.8	13.8
	ms_{22}	22.4	11.1	9.8	10.2	9.7	9.2
	ms_{23}	20.4	11.5	10.5	12.2	10.1	7.8
	ms_{24}	26.5	13.4	9.8	14.5	9.8	14.5
ts_3	ms_{31}	19.4	15.6	15.2	12.1	10.9	8.6
	ms_{32}	24.8	9.6	15.2	9.4	10.3	8.7
	ms_{33}	27.5	14.5	6.7	18.4	9.2	9.5
ts_4	ms_{41}	26.5	13.6	16.5	7.5	14.2	6.4
	ms_{42}	25.3	18.5	14.5	10.2	11.2	9.8
ts_5	ms_{51}	35	15.3	7.1	11.2	16.5	9.8
	ms_{52}	25.5	14.2	12.3	14.2	12.2	9.8
	ms_{53}	27.8	15.8	10.7	13	9.9	11
	ms_{54}	25.6	10.2	17.8	9.8	16.3	7.5
ts_6	ms_{61}	26.4	12.1	14.2	9.4	13.4	9.8
	ms_{62}	32.1	10.3	10.2	15.2	14.5	7.8
	ms_{63}	22.6	9.2	15.4	18.2	10.2	7.8
$ts7$	ms_{71}	24.5	10.5	14.4	9.4	11.2	10.5
	ms_{72}	28	9.8	15.4	16.4	9.2	9.6
ts_8	ms_{81}	22.2	13.4	18.8	12.3	11.8	9.6
	ms_{82}	27.7	7.4	13.1	16.3	13.8	6.4
	ms_{83}	22.4	6.9	22.3	7.8	14.8	6.5
	ms_{84}	27.6	16.6	14.8	16.5	7.9	11.2

The coordination data among manufacturing units were generated through random digits, as shown in Table 2. Then the disjunction graph model becomes a combinatorial optimization model, and it has met the coordination degree of the connection arc of each manufacturing unit. Besides, the manufacturing unit is selected by the following method, that is, One of the three designs in the manufacturing unit set 1 is selected, one of the four designs in the manufacturing unit set 1 is selected, and one of the three designs in the manufacturing unit set 3 is selected, one of the two designs in the manufacturing unit set 4 is selected, one of the four designs in the manufacturing unit set 5 is selected, and one of the three designs in the manufacturing unit set 6 is

selected, One of the two designs in the manufacturing unit set 7 is selected, and one of the four designs in the manufacturing unit set 8 is selected. Furthermore, each manufacturing unit has its own characteristics, and which has reflected in cost, efficiency, quality, time, agility, and greenness, besides, its degree of coordination between units is different as well.

Table 2 Degree of coordination between units

	ms_{21}	ms_{22}	ms_{23}	ms_{24}
ms_{11}	0.81	0.8	0.79	0.91
ms_{12}	0.65	0.85	0.65	0.56
ms_{13}	0.65	0.45	0.83	0.81

	ms_{41}	ms_{42}
ms_{11}	0.45	0.56
ms_{12}	0.64	0.65
ms_{13}	0.53	0.68

	ms_{31}	ms_{32}	ms_{33}
ms_{21}	0.91	0.68	0.73
ms_{22}	0.89	0.75	0.64
ms_{23}	0.79	0.72	0.86
ms_{24}	0.95	0.86	0.56

	ms_{51}	ms_{52}	ms_{53}	ms_{54}
ms_{31}	0.88	0.76	0.87	0.89
ms_{32}	0.66	0.79	0.88	0.92
ms_{33}	0.68	0.69	0.86	0.84

	ms_{51}	ms_{52}	ms_{53}	ms_{54}
ms_{41}	0.65	0.67	0.71	0.69
ms_{42}	0.95	0.66	0.86	0.83

	ms_{61}	ms_{62}	ms_{63}
ms_{51}	0.65	0.87	0.56
ms_{52}	0.86	0.89	0.64
ms_{53}	0.63	0.91	0.92
ms_{54}	0.65	0.75	0.45

	ms_{71}	ms_{72}
ms_{51}	0.81	0.78
ms_{52}	0.79	0.72
ms_{53}	0.67	0.86
ms_{54}	0.69	0.72

	ms_{81}	ms_{82}	ms_{83}	ms_{84}
ms_{61}	0.64	0.56	0.62	0.53
ms_{62}	0.81	0.72	0.65	0.58
ms_{63}	0.82	0.78	0.8	0.81

	ms_{81}	ms_{82}	ms_{83}	ms_{74}
ms_{71}	0.91	0.75	0.61	0.85
ms_{72}	0.89	0.78	0.89	0.87

5.2 The results

In this example, on the case of the discrete manufacturing system function objective decision model and algorithm (ACO-PSO), combined with MATLAB 7.0 simulation tool programming, the set of manufacturing units and the possible set of link arcs in the table are optimized and solved to obtain the functional objective combination of discrete manufacturing system. Parameter setting for programming solution was as follows. Particle swarm optimization part: population size1 = 20, $\omega = 1$, $c_1 = c_2 = 2$, the number of iterations 500. Ant colony algorithm part: ants size2 = 10，$\alpha = 1$，$\beta = 1$，$\rho = 0$，$Q = 1$, the number of iterations 60. The whole process is realized on the computer of Intel (R) Core (TM) i3 CPU550 @ 3.2 GHz ZN-2, RAM 2.O GB and Windows XP operating system. In the solving process, when the customer and the manufacturer have different requirements for the manufacturing task, which mainly reflected in the weight requirements for each objective, the decision-making results of the discrete manufacturing system functional objectives, which reflect the specific requirements of the combination optimization scheme. Decision-making schemes shown in Table 3 can be separately calculated to meet different requirements of the customer and the producer when they have own demand for each objective in the evaluation of the functional objectives. Taking the uniform weight as an example, each objective weight is $1/7$, in which the fitness function value is the total objective value, that is $Z^{"}$, meanwhile, the discrete manufacturing system function objective decision-making scheme is shown in Fig. 5.

As shown in Fig. 6, when PSO algorithm was employed alone, the initial convergence of the algorithm was faster. While with the number of iterations increased, the performance of local search ability was insufficient, and the convergence curve tended to level gradually, and which had resulted in that the solution was easy to fall into local optimum. Similarly, when the ACO algorithm was used alone, although the quality of the solution was higher than PSO, and it was also better at searching for the optimal solution. But its initial pheromone was relatively scarce, and which had resulted that its initial search is blind, time-consuming and slower. In short, this paper has designed a PSO-ACO algorithm for the functional decision making of discrete manufacturing systems, which fully circumvented the shortcomings of PSO and ACO algorithms and combined their advantages. Besides, it has greatly improved the accuracy and speed of the problem solved. Furthermore, both the GA algorithm and the proposed algorithm had the same minimum fitness value, and they could also avoid the premature and stagnant phenomenon of the solving process. However, obviously the GA algorithm took more time.

Table 3 Functional objective decision making of discrete manufacturing system under different weights

Objective weight							Fitness	Manufacturing unit combination
Cost	Efficiency	Quality	Time	Agility	Green	Coordination		
1/7	1/7	1/7	1/7	1/7	1/7	1/7	0.265	$ms_{12}ms_{22}ms_{32}ms_{41}ms_{52}ms_{63}ms_{84}$
0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.312	$ms_{14}ms_{22}ms_{33}ms_{42}ms_{53}ms_{62}ms_{82}$
0.3	0.2	0.1	0.1	0.1	0.1	0.1	0.327	$ms_{11}ms_{23}ms_{33}ms_{42}ms_{54}ms_{72}ms_{81}$
0.4	0.1	0.1	0.1	0.1	0.1	0.1	0.295	$ms_{13}ms_{22}ms_{32}ms_{42}ms_{53}ms_{62}ms_{84}$
0.1	0.1	0.2	0.2	0.2	0.1	0.1	0.276	$ms_{14}ms_{22}ms_{31}ms_{42}ms_{51}ms_{72}ms_{84}$
0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.267	$ms_{12}ms_{21}ms_{32}ms_{41}ms_{53}ms_{72}ms_{82}$
0.1	0.1	0.1	0.1	0.1	0.1	0.4	0.253	$ms_{14}ms_{24}ms_{32}ms_{42}ms_{53}ms_{63}ms_{83}$

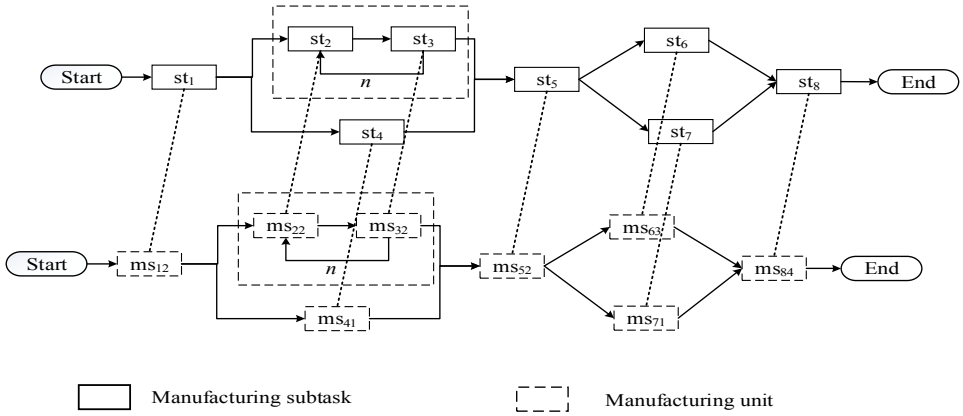


Fig. 5 Function objective decision making of discrete manufacturing system under uniform weight

In addition, the initial data at different scales are solved respectively. And the results have shown that, when the number of link arcs between the alternatives in the manufacturing unit set and the unit increased, the performance of the proposed algorithm was better.

As shown in Table 4 that the PSO-ACO algorithm proposed is the most accurate solution to the minimum of the objective function, followed by the ACO algorithm and GA algorithm, and the PSO algorithm is the worst. Besides, by analyzing the solutions obtained by running the three algorithms repeatedly, we found that the optimal solution obtained by PSO-ACO algorithm proposed has been floating around 0.387, and finally stabilized at 0.387. However, the results of the other two algorithms floated greatly and were not better than 0.387, which has shown that when PSO and ACO algorithms are used alone, their solutions tended to be limited due to their shortcomings and defects. Furthermore, it also indicated that the PSO-ACO algorithm had good solution performance.

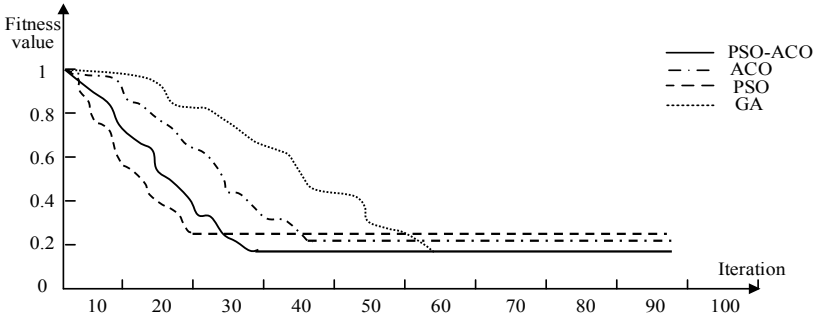


Fig. 6 Algorithm convergence curve

Table 4 Comparison of the results of four algorithms

Algorithm	Optimal solution objective function value	Optimum solution
PSOACO	0.387	$ms_{11}ms_{21}ms_{32}ms_{41}$ $ms_{53}ms_{62}ms_{83}$
ACO	0.418	$ms_{12}ms_{21}ms_{32}ms_{42}$ $ms_{53}ms_{71}ms_{83}$
PSO	0.463	$ms_{11}ms_{22}ms_{32}ms_{41}$ $ms_{52}ms_{71}ms_{81}$
GA	0.387	$ms_{11}ms_{21}ms_{32}ms_{41}$ $ms_{53}ms_{71}ms_{83}$

As shown in Table 5, the PSO-ACO algorithm had high search efficiency for three indexes on average value, worst value and average time, and these indicators were obtained from 30 consecutive runs. So it could be concluded that PSO-ACO algorithm is feasible and effective in discrete manufacturing system function objective decision-making.

In the functional objective decision of discrete manufacturing system shown in Table 6, the one-way optimal combination schemes has been given respectively, which was in cost, efficiency, quality, time, agility, greenness and coordination, meanwhile, and the overall optimal decision scheme was given in the last row, as shown in Table 6. Therefore, the decision-making scheme can be described as Fig. 7.

Table 5 Comparison of four algorithms

Algorithm	Optimal value	Average value	Worst value	Average time
PSACO	0.372	0.407	0.498	32
ACO	0.421	0.492	0.585	40
PSO	0.477	0.533	0.614	27
GA	0.374	0.451	0.570	43

Table 6 The best solution set of different objectives in discrete manufacturing system

Objective	Set of Solutions	Cost	Productivity	Quality	Time	Agility	Green	Coordination	Comprehensive satisfaction
Low cost	$ms_{11}ms_{22}$ $ms_{32}ms_{41}$ $ms_{52}ms_{63}$ ms_{84}	172.8	84.3	97.4	93.5	73.4	60.2	0.51	0.580
High productivity	$ms_{12}ms_{22}$ $ms_{31}ms_{42}$ $ms_{54}ms_{62}$ ms_{82}	178.1	85.6	97.1	85.3	85.3	63.4	0.54	0.582
High quality	$ms_{11}ms_{24}$ $ms_{33}ms_{41}$ ms_{54} $ms_{72}ms_{81}$	179.7	84.9	98.4	96.4	79.4	64.2	0.55	0.593
Short time	$ms_{13}ms_{22}$ $ms_{33}ms_{42}$ $ms_{52}ms_{62}$ ms_{83}	179.7	85.3	97.3	84.7	86.1	59.4	0.57	0.607
Great agility	$ms_{13}ms_{22}$ $ms_{33}ms_{42}$ $ms_{51}ms_{63}$ ms_{81}	183.4	85.2	98.0	87.2	89.2	58.5	0.58	0.602
Good greenness	$ms_{12}ms_{22}$ $ms_{33}ms_{42}$ ms_{54} $ms_{72}ms_{83}$	176.8	86.5	103	85.3	79.3	66.2	0.61	0.558
coordination	$ms_{12}ms_{24}$ $ms_{31}ms_{42}$ ms_{52} $ms_{72}ms_{83}$	181.5	84.6	98.1	88.7	80.3	57.9	0.65	0.589
Comprehensive satisfaction	$ms_{11}ms_{21}$ $ms_{32}ms_{41}$ $ms_{53}ms_{62}$ ms_{83}	182.4	84.7	98.1	85.6	82.4	61.3	0.62	0.613

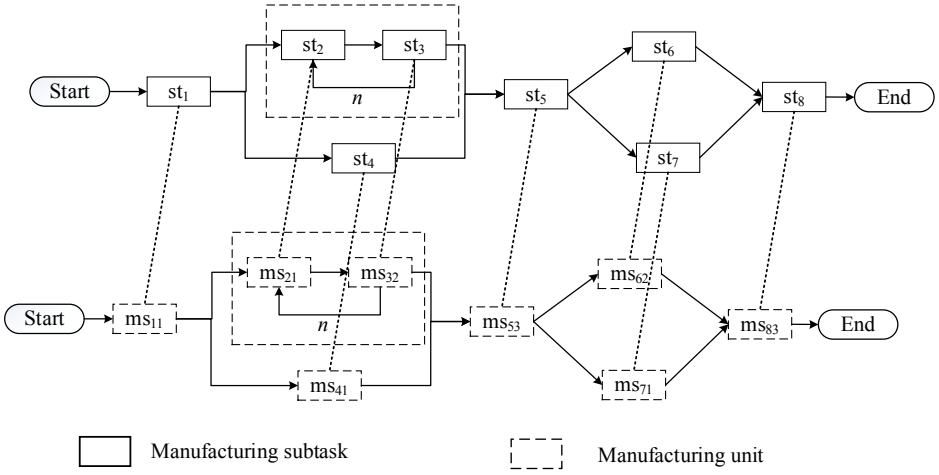


Fig. 7 The overall optimal solution of discrete manufacturing system functional objectives

This method could satisfy the preference of the functional objectives of discrete manufacturing system for different sub-tasks. Besides, when using the functional objective optimization combination method and PSO-ACO method, we could obtain the Pareto optimal state solution set of the discrete manufacturing system function objective decision and the optimal scheme in a specific aspect. At the same time, the optimal synthesis scheme could be given, which was helpful for decision-makers and customers to choose the optimal scheme and improve the environmental adaptability of manufacturing system.

6. Conclusion

This paper has established a decision model, and in which manufacturing unit was considered as a discrete manufacturing system function objective carrier. And it had different characteristics, such as cost, efficiency, quality, time, agility and green, and the coordination degree between units was also fluctuating. Besides, the decision-making model, production capacity, resource constraints and objective thresholds have constituted to a decision-making system for functional objectives of discrete manufacturing systems, and it also has quantified the data. The PSO-ACO method is used to solve the problem, and the preference of the combination of customer weight and enterprise weight was given as the combination weight. Finally, an example was given to verify the proposed method, which also proved the practicability of the algorithm.

Acknowledgement

This work was financially supported by Liaoning Province Planning Office of Philosophy and Social Science (L18BJY029), Shenyang Planning Office of Philosophy and Social Science (18ZX019), the Liaoning Province Department of Education Project (WGD2016002), Shenyang Science and Technology Innovation Knowledge Base (SYKJ201807,201806), and Shenyang Federation Social Science Circles (Grant no. SYSK2018-05-05), the Key Program of Social Science Foundation of Liaoning Province (Grant No. L15AGL013), and the Natural Science Foundation of Liaoning Province (Grant No. 201602545). The authors wish to acknowledge the contribution of Liaoning Key Lab of Equipment Manufacturing Engineering Management, Liaoning Research Base of Equipment Manufacturing Development, Liaoning Key Research Base of Humanities and Social Sciences, Research Center of Micromanagement Theory, and Shenyang Association for Science and Technology.

References

[1] Wang, Y.-H., Lee, C.-H., Trappey, A.J.C. (2017). Service design blueprint approach incorporating TRIZ and service QFD for a meal ordering system: A case study, *Computers & Industrial Engineering*, Vol. 107, 388-400, doi: 10.1016/j.cie.2017.01.013.

[2] Francia, D., Caligiana, G., Liverani, A., Frizziero, L., Donnici, G. (2017). PrinterCAD: A QFD and TRIZ integrated design solution for large size open moulding manufacturing, *International Journal on Interactive Design and Manufacturing*, Vol. 12, No. 1, 81-94, doi: 10.1007/s12008-017-0375-2.

[3] Oh, S., Cho, B., Kim, D.-J. (2017). Development of an exportable modular building system by integrating quality function deployment and TRIZ method, *Journal of Asian Architecture and Building Engineering*, Vol. 16, No. 3, 535-542, doi: 10.3130/jaabe.16.535.
[4] Rami, M.A., Napp, D. (2016). Discrete-time positive periodic systems with state and control constraints, *IEEE Transactions on Automatic Control,* Vol. 61, No. 1, 234-239, doi: 10.1109/TAC.2015.2438428.
[5] Fernández, J., Tori, C., Zuccalli, M. (2016). Lagrangian reduction of discrete mechanical systems by stages, *Journal of Geometric Mechanics*, Vol. 8, No. 1, 35-70, doi: 10.3934/jgm.2016.8.35.
[6] Vieira, A.D., Santos, E.A.P., de Queiroz, M.H., Leal, A.B., de Paula Neto, A.D., Cury, J.E.R. (2017). A method for PLC implementation of supervisory control of discrete event systems, *IEEE Transactions on Control Systems Technology*, Vol. 25, No. 1, 175-191, doi: 10.1109/TCST.2016.2544702.
[7] Sopasakis, P., Sarimveis, H. (2017). Stabilising model predictive control for discrete-time fractional-order systems, *Automatica*, Vol. 75, 24-31, doi: 10.1016/j.automatica.2016.09.014.
[8] Alicandro, R., Braides, A., Cicalese, M. (2006). Phase and anti-phase boundaries in binary discrete systems: A variational viewpoint, *Networks & Heterogeneous Media*, Vol. 1, No. 1, 85-107, doi: 10.3934/nhm.2006.1.85.
[9] Feniser, C., Burz, G., Mocan, M., Ivascu, L., Gherhes, V., Otel, C.C. (2017). The evaluation and application of the TRIZ method for increasing eco-innovative levels in SMEs, *Sustainability*, Vol. 9, No. 7, 1125, doi: 10.3390/su9071125.
[10] Hsieh, H.-N., Chen, J.-F., Do, Q.H. (2017). A creative research based on DANP and TRIZ for an innovative cover shape design of machine tools, *Journal of Engineering Design*, Vol. 28, No. 2, 77-99, doi: 10.1080/09544828.2016.1272100.
[11] Hone, A.N.W., Petrera, M. (2009). Three-dimensional discrete systems of Hirota-Kimura type and deformed Lie-Poisson algebras, *Journal of Geometric Mechanics, Vol.* 1, No. 1, 55-85, doi: 10.3934/jgm.2009.1.55.
[12] Mahmoud, M.S., Shi, P. (2016). Optimal guaranteed cost filtering for Markovian jump discrete-time systems, *Mathematical Problems in Engineering*, Vol. 2004, No. 1, 33-48, doi: 10.1155/S1024123X04108016.
[13] Tang, M., Gong, D., Liu, S., Lu, X. (2017). Finding key factors affecting the loactions of electric vehicle charging stations: A simulation and ANOVA approach, *International Journal of Simulation Modelling*, Vol. 16, No. 3, 541-554, doi: 10.2507/IJSIMM16(3)CO15.

APEM journal
Advances in Production Engineering & Management
Volume 13 | Number 4 | December 2018 | pp 405–416
https://doi.org/10.14743/apem2018.4.299
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Effect of process parameters on the surface roughness of aluminum alloy AA 6061-T6 sheets in frictional stir incremental forming

Azpen, Q.[a,*], **Baharudin, H.**[b], **Sulaiman, S.**[c], **Mustapha, F.**[d]

[a,b,c]Universiti Putra Malaysia, Faculty of Engineering, Department of Mechanical and Manufacturing Engineering, Serdang, Malaysia
[d]Universiti Putra Malaysia, Faculty of Engineering, Department of Aerospace Engineering, Serdang, Malaysia
[a]Middle Technical University, Institute of Technology, Baghdad, Iraq

ABSTRACT

Incremental Sheet Forming (ISF) is characterized by essential flexibility, great formability, and low forming forces and cost compared to the conventional sheet metal forming processes. ISF was born as an advance sheet metal forming process to perfectly fit previous requirements. Nevertheless, growing demand to apply the lightweight materials in several fields was placed this developed process in a critical challenge to manufacture the materials with unsatisfied formability especially at room temperature. Thus, utilizing the heat at warm and hot condition in some ISF processes has been introduced to solve this problem. Among all heat-assisted ISF processes, frictional stir-assisted Single Point Incremental Forming (SPIF) was presented to deal with these materials. In this work, this emerging process was utilized to manufacturing products from AA6061-T6 aluminum alloy. Experimental tests were performed to study the influence of main parameters like tool rotation speed, feed rate, step size and tool size on the surface roughness of the produced parts. A Taguchi method and varying wall angle conical frustum (VWACF) test were used in the present work. The results find that tool diameter has a significant impact on the internal surface roughness produced via the forming process with a percentage contribution of 93.86 %. The minimum value of the surface roughness was 0.3 μm.

ARTICLE INFO

Keywords:
Friction stir forming;
Incremental sheet forming (ISF);
Heat-assisted ISF;
Surface roughness;
Aluminum alloy (AA6061-T6)

**Corresponding author:*
qasimmhalhal@gmail.com
(Azpen, Q.)

Article history:
Received 11 July 2018
Revised 26 August 2018
Accepted 28 August 2018

1. Introduction

Currently, there is a growing market in the manufacturing of customized, rapid prototyping and low-cost sheet parts with small to medium batches (particularly in transportation, artificial medical alternatives, and aerospace industries) [1, 2]. The main reason for employing tool rotational speed in SPIF is to improve the formability of lightweight and hard-to-form materials which characterized by low formability at room temperature [3-5]. In addition, it leads to decrease the forces via the forming process [6-8]. Indeed, there are some drawbacks and tradeoffs that influence as a result of employing the spindle speed in SPIF. Low surface quality, lubricant failure, and high tool wear rate are the disadvantages of frictional stir incremental forming process [9]. In addition, SPIF at high rotation speed promotes the probability for developing tool marks on the worked sheets [10].

The surface finish or surface roughness is a serious drawback in ISF, which limits the expansion of this process in different applications. To obtain a better surface texture, it is important to control several processes and material factors like forming angle, tool rotation, tool size and shape, step size, sheet thickness, and friction and lubricant. Thus, the researchers considered the influences of these main factors on the final surface topography of the produced parts in SPIF.
A study conducted by Durante *et al.* [6] aimed to investigate the influence of the tool rotational speeds and its directions on the surface texture of aluminum alloy AA7075-T0 formed by SPIF. The experimental results proved that no significant effect of these two parameters was present in the studied speed range between 0-600 rpm, and the obtained varied values of the surface roughness were mainly dependent on whether the tool was rotated or not.

During SPIF of the aluminum alloy AA3003-H14, a model was established by Hamilton and Jeswiet [11] which can be employed to improve the external surface of manufacturing SPIF parts by selecting adequate forming parameters via a process such as feed rate and tool rotation speed at high speeds. This presented model can predict the orange peel effect and provide a good guide to enhance the surface quality. In addition, the surface roughness for the parts with high rotation speed/feed rate is less than those of with a low ratio.

Good surface roughness results were obtained during the manufacturing of medical parts by SPIF from the known titanium alloy Ti-6Al-4V by Olesksik *et al.* [12]. The obtained surface finish of the formed parts were influenced by the forming tool roughness and friction case at the tool-sheet zone.

In fact, the final formed angle in SPIF is used as an index for both formability and surface roughness where the change in stretching value in the formed part leads to the change in both the forming angle and surface finish. In this way, Bhattacharya *et al.* [13] investigated the impact of tool size (4 mm, 6 mm , and 8 mm), step size (0.2 mm, 0.8 mm, and 1 mm), and wall angle (20°, 40°, and 60°) on the surface roughness of aluminum alloy AA5052 via SPIF. The results of experiments showed that the surface quality of the formed parts decreases as to the increase in tool sizes for all step sizes. In addition, surface finish decreases due to the increasing of the forming angle.

Palumbo and Brandizzi [14] proved that both the surface roughness and the part's accuracy are influenced by the tool rotation speed when the forming of the titanium alloy Ti6Al4V was studied. The spindle rotation range of 800-1600 rpm was with two values of step sizes; 0.5 mm and 1 mm. The value of R_a became 011.9 μm compared to the initial sheet roughness of 0.5 μm.
Ambrogio *et al.* [15] performed an experimental study on three aluminum alloys, AA1050-0, AA5754, and AA6082-T6, with different sheet thicknesses. It was proven that the step size, forming angle, and sheet thickness have a significant impact on the surface roughness of the shaped parts, while it had an insignificant impact on the feed rate.

In this regard, Silva *et al.* [16] studied the influence of both the step size and feed rate on the surface roughness of SAE 1008 steel material. It was shown that an adequate roughness could be obtained with a feed rate and step size of 8400 mm/min and 0.2 mm, respectively. Lasunon *et al.* [17] examined the effect of some factors on the surface finish. Their results proved that the forming angle, step size, and its interaction affected the achieved surface texture, while there was little influence on the feed rate.

The good surface finish can be achieved depending on the tool trajectory. Usually, the tool trajectory with a constant step depth leaves marks at the end of each circle of the path and, therefore, produces a poor surface quality; especially with high step values compared to the spiral tool path [18]. Skjoedt *et al.* [19] proved that scarring can be removed by using a spiral trajectory during SPIF. Lu *et al.* [20] given a tool path algorithm based on specified critical edges. A superior surface roughness can be obtained by using this algorithm with respect to the traditional tool path employed in ISF.

An empirical research was conducted by Liu *et al.* [21] to investigate the influence of tool size, feed rate, step size, and sheet thickness on the surface texture of the final part made from aluminum alloy AA7075-T0. The response surface methodology and Box-Behnken design were applied to analyze the results. Better surface roughness was achieved with parameter values of 25

mm for tool size, 6000 mm/min for feed rate, 0.39 mm for step size, and 1.6 mm for sheet thickness.

Mugendiran *et al*. [22] built a quadratic model with second order based on three process parameters (tool rotation, feed rate, and tool diameter) to estimate the influence of the mentioned variables on both the surface finish and wall thickness distribution during the forming of aluminum alloy AA5052. Optimum values of surface roughness R_a and final sheet thickness (t) were 2.45 μm and 0.753 mm, respectively. These optimal values were obtained at rotation speed, feed rate, and step size of 1931 rpm, 654 mm/rev, and 0.65 mm, respectively.

Another study was conducted by Lu *et al.* [23] to determine the impact of the tool design on the surface quality of four aluminum alloys named AA6111, AA5052, AA2024, and AA1100. The obtained results concluded that better surface roughness could be achieved with new oblique roller-ball tool (ORB) rather than the conventional tool. The employment of ORB helped in reducing the friction at the tool-sheet zone and at the same time, reduced the forming loads and increased the formability of the studied materials.

A detailed experimental study by Azevedo *et al.* [24] aimed to estimate the effect of several types of lubricants on the surface roughness for steel DP780 and aluminum alloy AA1050-T4. It was concluded that the existence of lubricant is an important factor to obtain better surface texture. This finding supported the results of previous studies [25-27].

In this work, friction stir-assisted SPIF was utilized to manufacturing AA6061-T6 sheets that have been utilized in several applications in industrial sectors. Besides the mentioned benefits, friction stir-assisted SPIF shows superior profits, where, it does not need an exterior heating source and the surface finish is better than the other two heat-assisted ISF types: electric-assisted ISF and laser-assisted ISF.

2. Materials and methods

2.1 Material

Uniaxial tensile test was achieved to get the stress-strain curve of AA 6060-T6 sheet with a thickness of 2 mm. Fig. 1 presents the specimen dimensions which are according to ASTM E8M standard.

Fig. 2 and Table 1 describe the true stress-strain curve and the chemical composition of the material used, respectively. It is clear that the material has a suitable total strain at fracture, which is preferred in incremental sheet metal forming.

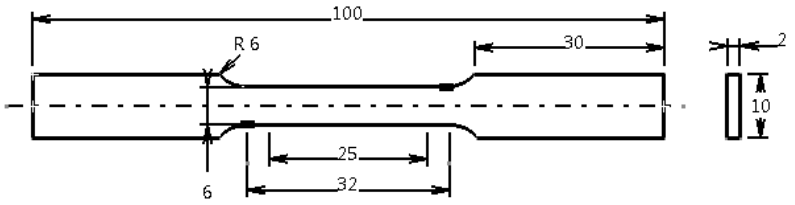


Fig. 1 Specimen dimensions of the uniaxial tensile test (dimensions in mm)

Table 1 Chemical composition (wt %) of the material

Material	Si	Fe	Cu	Mn	Mg	Cr	Ni	Zn	Ti	Al
AA6061-T6	0.52	0.19	0.27	0.07	0.91	0.1	-	0.02	0.01	97.91

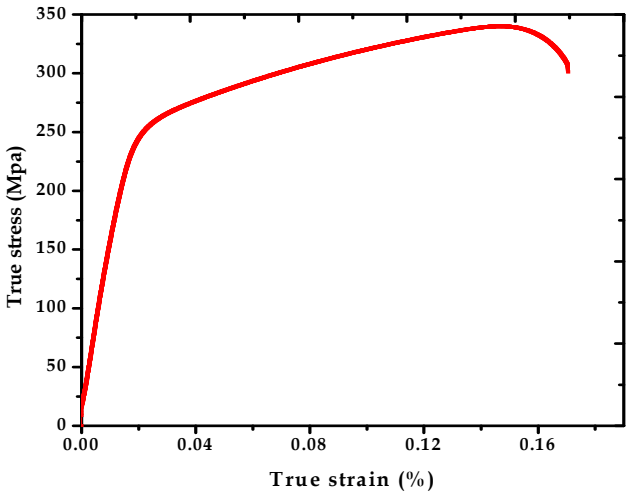


Fig. 2 The stress-strain curve of AA6061-T6

2.2 Experimental setup

The necessary task of the jig, which use in the forming process, is tightly hold the sheet specimen with both clamping and backing plates. Forming jig included of four clamping plates, backing plate, four columns and base plate. The dimensions of the backing plate are 170 × 170 × 20 mm with a central hole of 70 mm in diameter, which represent the outer diameter of final product. In order to get a smooth material forming, the inner diameter of the backing plate was filleted with 60 mm radius. On the other side, the aluminum sheet is with dimensions of 150 × 150 × 2 mm. The whole jig assemble was mounted to the bed of CNC milling machine (OKUMA MX 45VA). Fig. 3 displays the experimental setup of the forming jig.

Two forming tools with hemispherical ends were employed in the experimental tasks. The tools are with two different diameters, 10 mm and 15 mm and with a same of length of 110 mm. Moreover, these tools were hardened and tempered with 60 HRC and made from high-speed steel (HSS) material. In order to decrease friction effect at the contact zone thereby increase the tools life and surface quality of the final product, tool tips were polished. Fig. 4 explains the dimensions of the tools used in the experiments.

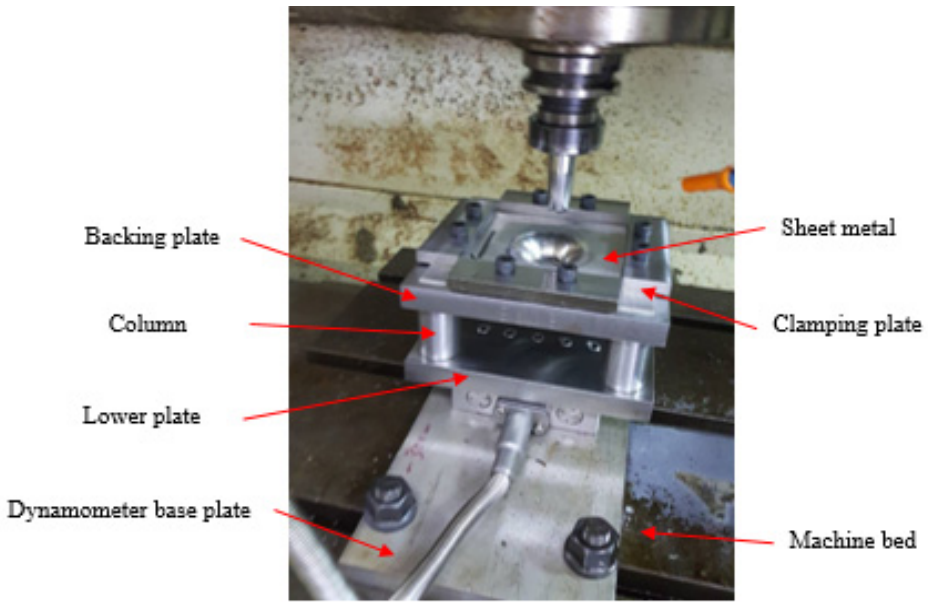


Fig. 3 The forming jig

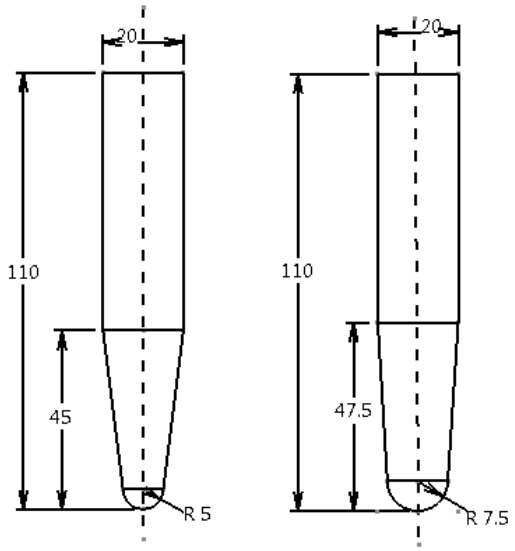


Fig. 4 The forming tools (all dimensions in mm)

2.3 Experiments

A varying wall angle conical frustum test (VWACF) was used to achieve the tests because of its homogenous geometry with the symmetrical parts [38]. The intended model of the product was designed to get maximum diameters (outer and inner of 70 mm and 12 mm, respectively), a height of 41 mm and a radius with 60 mm of the varying slops. Fig. 5 explains the designed dimensions of the targeted cone.

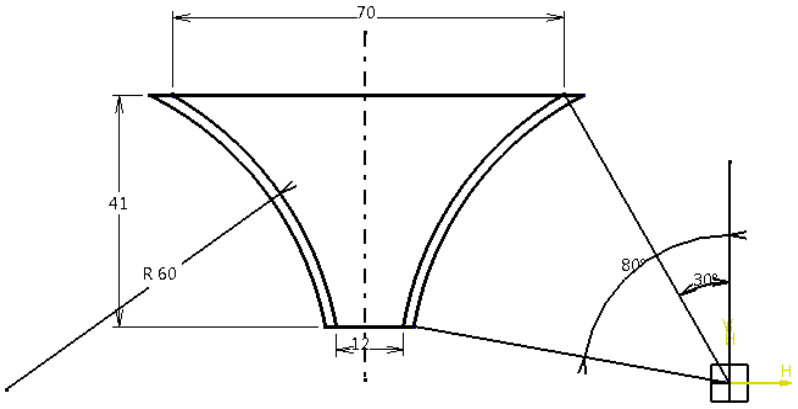


Fig. 5 The conical profile (dimension in mm)

A spiral trajectory of the forming tool with a certain step size was designed to generate the tool path. This path can be characterized by a pure stretch deformation during the forming process, which helps to create a sheet thickness that uniformly distributed [39]. Moreover, it assists to remove the peaks of the forming forces and at the same time, no stretch marks can leave on the working sheet surface. On the other hand, these sockets regularly happen with counter type. The CAD/CAM was used to create the product profile and generate the spiral tool path by NC code, as displayed in Fig. 6.

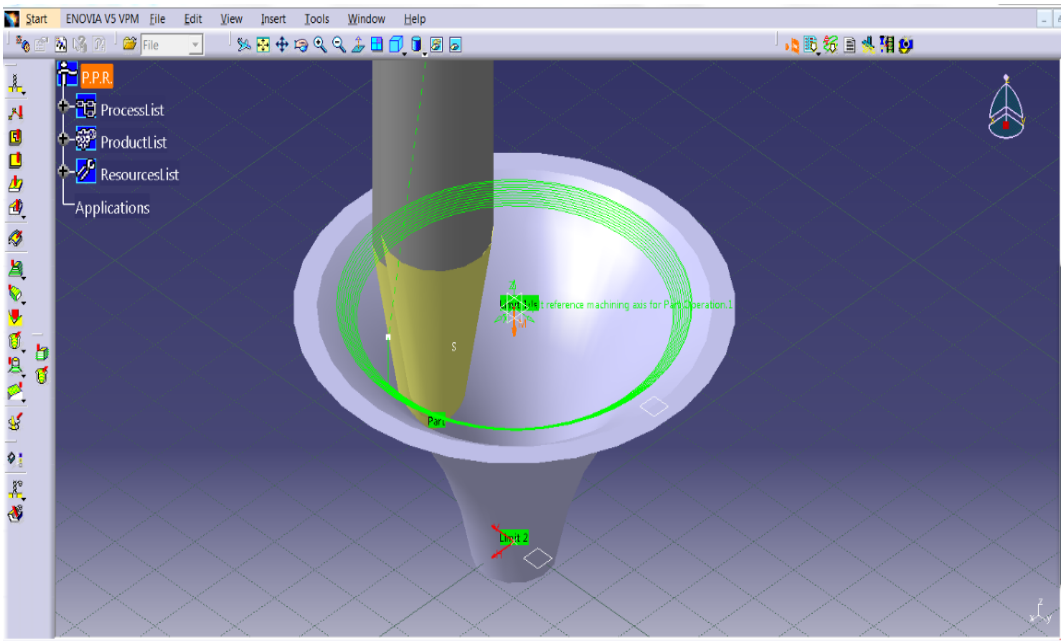
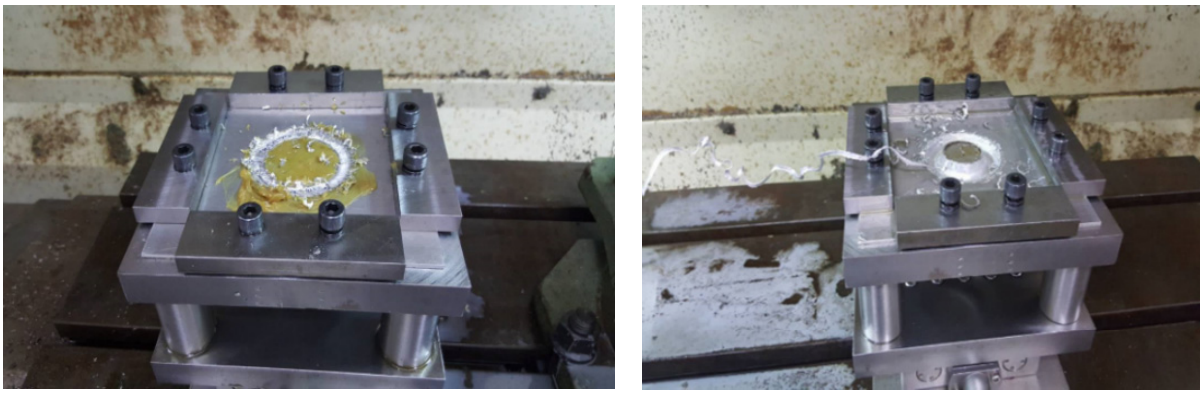


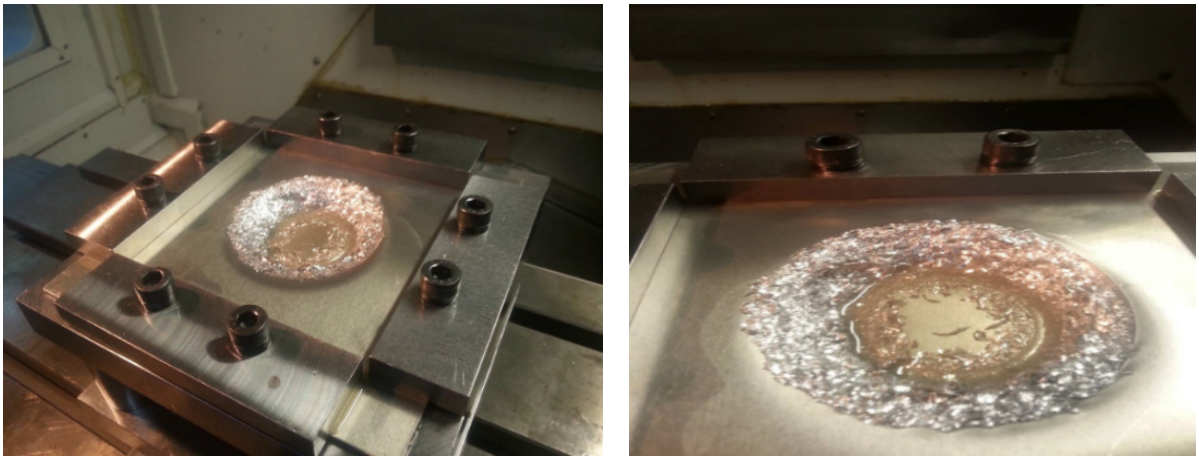
Fig. 6 Generating the tool path by CATIA

The continual motion of the forming tool via forming process leads to a local heating at the contact zone due to the local friction. In addition, this heating increases the rate of the tools wear. This will affect both surface roughness and geometric accuracy of the produced parts. These harmful effects can be prevented by using different types of lubricants. In this study, lubricant SAE 0W-40 was employed to diminish the friction effects. Taguchi technique was employed to help in the design of the tests with a minimum number of runs to save the time and overall cost [40, 41]. Design of experiment (DoE) which comprises of selection process parameters and their influential levels that depended on the previous studies. From these studies, it was concluded these factors and their levels are extremely affected by the material properties. In order to find the correct and suitable process parameter levels that can be used to obtain successful sets of experiments, many primary trials were conducted. Fig. 7 (a) and (b), and Fig. 8 (a) and (b) show the first failed trails due to the use of high rotation speed and feed rate, and small tool size, respectively.



(a) (b)

Fig. 7 Samples failed due to use high levels of rotation speeds and feed rates



(a) (b)

Fig. 8 Samples failed due to use a small tool size

Table 2 Process factors and their levels

Description	Factor	Level 1	Level 2	Level 3	Level 4
Tool rotation speed (rpm)	ω	50	400	800	1200
Feed rate (mm/min)	f	250	500	-	-
Step size (mm)	z	0.2	0.5	-	-
Tool size (mm)	D	10	15	-	-

Table 3 Orthogonal array L8 (4^1. 2^3) of the experiments tests

Test	ω (rpm)	f (mm/min)	z (mm)	D (mm)
1	1	1	1	1
2	1	2	2	2
3	2	1	1	2
4	2	2	2	1
5	3	1	2	1
6	3	2	1	2
7	4	1	2	2
8	4	2	1	1

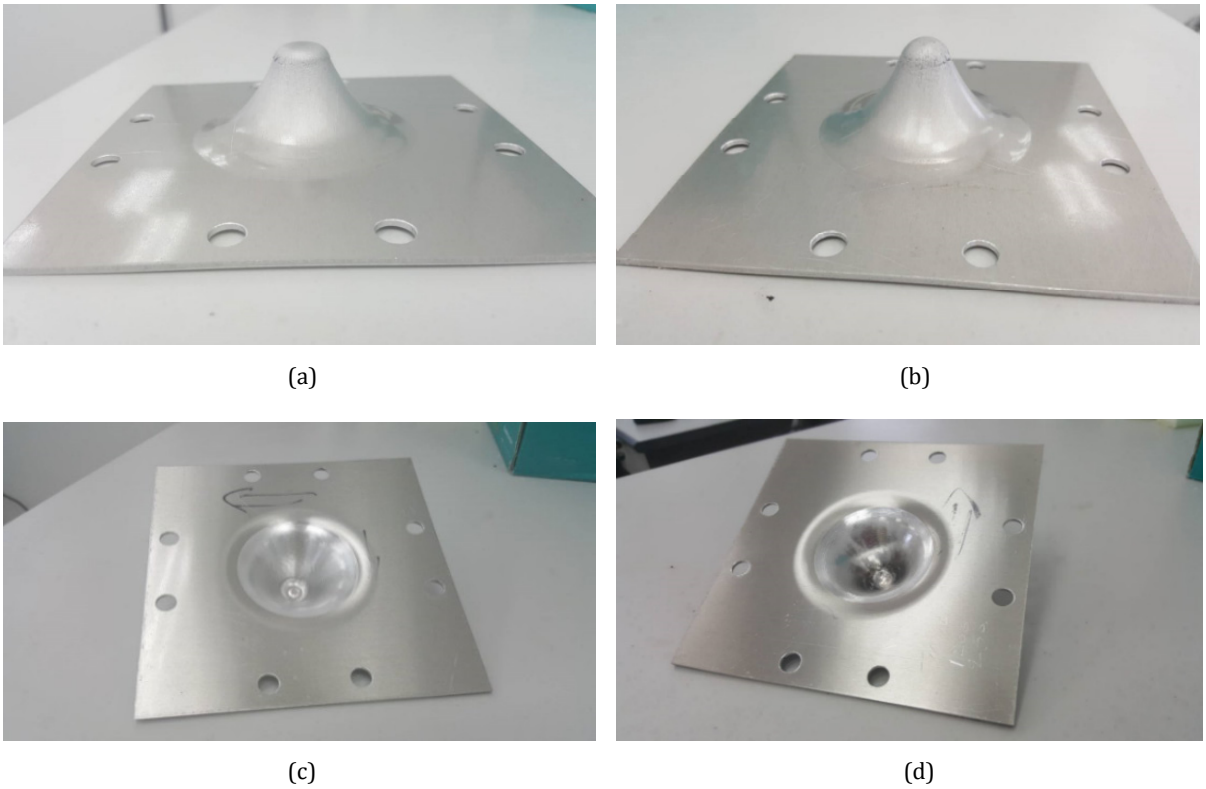
Tables 2 and 3 represent the process parameters, their levels, and the orthogonal array, respectively.

3. Results and discussion

A number of experimental tests were carry out to assess the effect of the tool rotation speed (ω), feed rate (f), step size (z) and tool size (D) on the final surface texture created through the SPIF. The experiments were stopped when the parts fracture. Where Fig. 9 (a), (b), (c) and (d) is demonstrated the samples that succeeded with the correct selection of parameter levels according to the mentioned designed array.

One of the main draw backs that accompany incremental sheet forming is the poor surface quality of the produced components [23]. Thus, appropriate combination and optimization of forming parameters is a challenge and an imperative issue to manufacture parts with excellent surface finish and other desirable process aspects; such as formability and forming forces. To achieve this goal, the Taguchi technique together with analysis of variance ANOVA, were employed to examine the influence of the tool rotation, feed rate, step size, and tool size on the obtained surface roughness. These four forming parameters have significant effects on SPIF, as mentioned in the literature.

The experimental results for the surface roughness R_a and the congruous S/N ratios are recorded in Table 4. Moreover, the surface roughness values for the AA6061-T6 sheets as received are 0.175 μm and 0.411 μm, with and across the rolling direction, respectively.



(a) (b)

(c) (d)

Fig. 9 Samples that succeeded with the correct selection of parameter levels

Table 4 The DoE matrix and the results for surface roughness and S/N ratios

Run	ω (rpm)	f (mm/min)	z (mm)	D (mm)	Across the forming tool path	
					Ra(μm)	S/N ratio
1	50	250	0.2	10	1.62	-4.1903
2	50	500	0.5	15	0.719	2.8654
3	400	250	0.2	15	0.581	4.7165
4	400	500	0.5	10	1.536	-3.7278
5	800	250	0.5	10	1.44	-3.1672
6	800	500	0.2	15	0.3	10.4576
7	1200	250	0.5	15	0.469	6.5765
8	1200	500	0.2	10	1.391	-2.8665

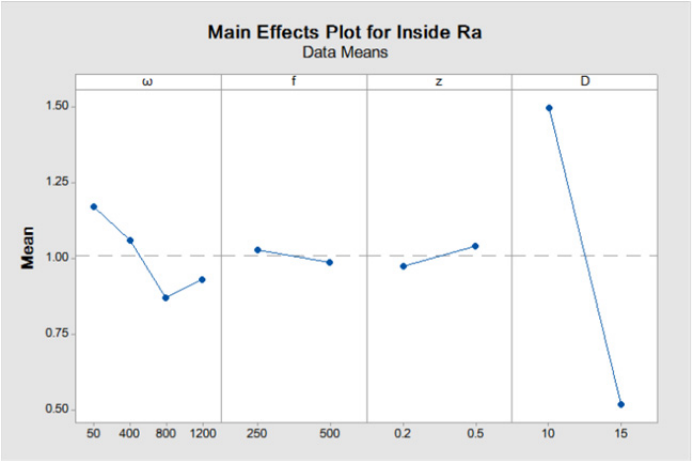


Fig. 10 The main effects of the various parameters on the surface finish

The main effect of the considered factors on the surface finish is presented Fig. 10. More or fewer impacts of the levels of these parameters on the output response can be noted. This graph shows the effect of tool size is a significant on the surface roughness. The other parameters such as rotation speed, feed rate, and step size have a less or negligible effect on the output.

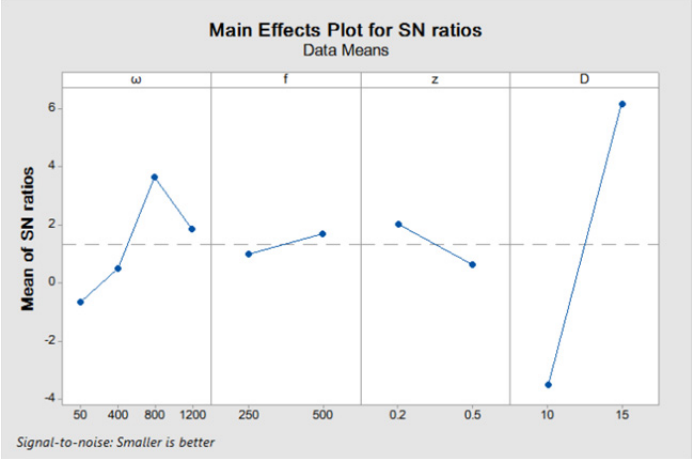


Fig. 11 Main effect for SN ratios for the surface roughness

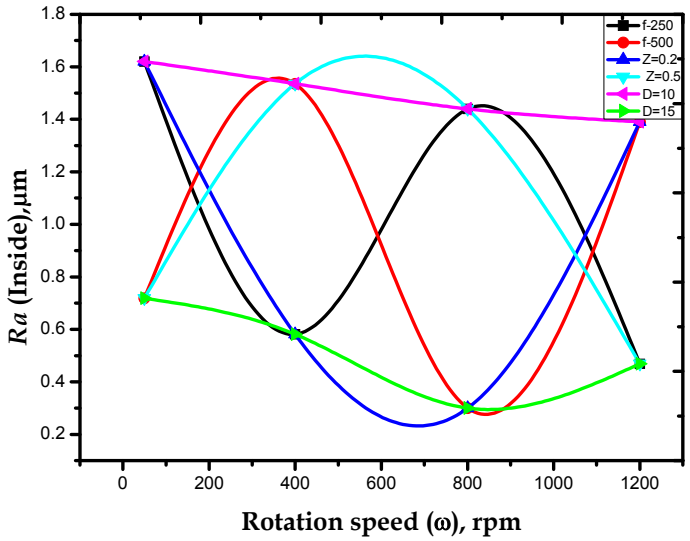


Fig. 12 The interaction effect of various factors on the surface roughness

Table 5 Analysis of variance for the surface roughness

Source	DF	Adj SS	Adj MS	F-Value	P-Value	Significant	Contribution (%)
Regression	4	2.01239	0.50310	47.35	0.005	Yes	
ω	1	0.08094	0.08094	7.62	0.070	-	3.96
f	1	0.00336	0.00336	0.32	0.613	-	0.16
z	1	0.00925	0.00925	0.87	0.420	-	0.45
D	1	1.91884	1.91884	180.59	0.001	Yes	93.86
Error	3	0.03188	0.01063				1.56
Total	7	2.04427					100

Table 6 Statistical results of the developed regression equation of the surfaces roughness

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	3.581	0.237	15.08	0.001	-
ω	-0.000234	0.000085	-2.76	0.070	1.00
f	-0.000164	0.000292	-0.56	0.613	1.00
z	0.227	0.243	0.93	0.420	1.00
D	-0.1959	0.0146	-13.44	0.001	1.00

The optimal condition of the surface roughness R_a of the AA6061-T6 sheet formed by frictional-stir assisted SPIF was determined by the Taguchi analysis. Consistent with this method, the greater the value of the S/N ratio is the superior the aggregate performance is. The case indicates that the parameter levels with the highest S/N ratio should be designated as the best levels. In this study, the optimal condition for the process parameters, which provided a minimum *Ra*, was within the run number 6, as shown in Table 4, and Figs. 11 and 12.

The analysis of variance ANOVA helped to create these relative impacts of parameters and their percentages contribution to the surface roughness, shown in Table 5 while Table 6 presents the coefficients of the regression equation.

This regression equation was established based on the experimental results of the surface roughness *Ra*; the (Eq.1) can describe it.

$$\text{Inside } Ra\ (\mu m) = 3.581 - 0.000234\ \omega - 0.000164\ f + 0.227\ z - 0.1959\ D \quad (1)$$

The fitting of the regression model is given by the determination coefficient R^2. The value of this coefficient refers to the close fitting of the regression equation. The values of the R^2, adjusted R^2, and predicted R^2 are 98.44 %, 96.36 %, and 87.86 %, respectively. Therefore, regarding the values of these coefficients, the established regression equation fits well and describes the surface roughness response. Lastly, normal distribution plot, Fig. 13, clarified that the residuals track the normal distribution. It can be noted that the regression equation has a good fit to their experimental data and are reliable to use.

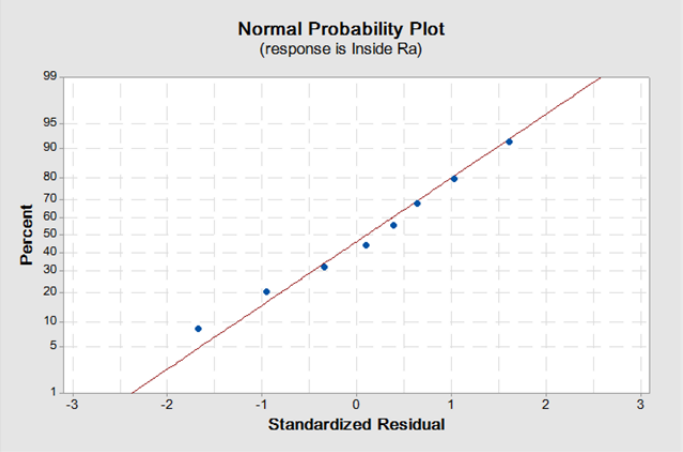


Fig. 13 Normal distribution of the surface roughness

4. Conclusion

In the present study, friction stir-assisted SPIF was performed to deform AA606-T6sheets. The purpose is to study the impact of certain process factors on the surface roughness of the produced parts. The results can be concluded in the following worthy points:

- The diameter of the forming tool have a significant impact on the internal surface roughness produced via the forming process of AA6061-T6. To attain an acceptable surface roughness, the percentage contribution of this parameter was 93.86 %.
- An optimal process parameters was achieved for the surface roughness during the forming process. The minimum value of the surface roughness was 0.3 µm at ω = 800 rpm, f = 500 mm/min, z = 0.2, and D = 15 mm.
- The value of the determination coefficient R^2 of the established regression equation of the surface roughness was 98.44 %. This high value refer to the close fitting of the suggested equation to describe the expected experimental data; it also means the response values highly adhere to the normal distribution.

Acknowledgments

The authors would like to thank Iraqi Government for the PhD scholarship and Research Management Centre, Universiti Putra Malaysia with the research grant (GP-IPS/2016/9479500), which enable the research to be carried out successfully.

References

[1] Ambrogio, G., Filice, L., Gagliardi, F. (2012). Formability of lightweight alloys by hot incremental sheet forming, *Materials & Design*, Vol. 34, 501-508, doi: 10.1016/j.matdes.2011.08.024.

[2] Bao, W., Chu, X., Lin, S., Gao, J. (2015). Experimental investigation on formability and microstructure of AZ31B alloy in electropulse-assisted incremental forming, *Materials & Design*, Vol. 87, 632-639, doi: 10.1016/j.matdes.2015.08.072.

[3] Nee, A.Y.C. (2015). *Handbook of manufacturing engineering and technology*, Springer, London, UK, doi: 10.1007/978-1-4471-4670-4.

[4] Xu, D., Lu, B., Cao, T., Chen, J., Long, H., Cao, J. (2014). A comparative study on process potentials for frictional stir-and electric hot-assisted incremental sheet forming, *Procedia Engineering*, Vol. 81, 2324-2329, doi: 10.1016/j.proeng.2014.10.328.

[5] Asghar, J., Reddy, N.V. (2013). Importance of tool configuration in incremental sheet metal forming of difficult to form materials using electro-plasticity, *In: World Congress on Engineering*, WCE 2013, London, United Kingdom.

[6] Durante, M., Formisano, A., Langella, A., Minutolo, F.M.C. (2009). The influence of tool rotation on an incremental forming process, *Journal of Materials Processing Technology*, Vol. 209, No. 9, 4621-4626, doi: 10.1016/j.jmatprotec.2008.11.028.

[7] Baharudin, B.T.H.T., Azpen, Q.M., Sulaima, S., Mustapha, F. (2017). Experimental investigation of forming forces in frictional stir incremental forming of aluminum alloy AA6061-T6, *Metals*, Vol. 7, No. 11, 484, doi: 10.3390/met7110484.

[8] Bagudanch, I., Centeno, G., Vallellano, C., Garcia-Romeu, M.L. (2013). Forming force in single point incremental forming under different bending conditions, *Procedia Engineering*, Vol. 63, 354-360, doi: 10.1016/j.proeng.2013.08.207.

[9] Hagan, E., Jeswiet, J. (2004). Analysis of surface roughness for parts formed by computer numerical controlled incremental forming, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture,* Vol. 218, No. 10, 1307-1312, doi: 10.1243/0954405042323559.

[10] Jeswiet, J., Micari, F., Hirt, G., Bramley, A., Duflou, J., Allwood, J. (2005). Asymmetric single point incremental forming of sheet metal, *CIRP Annals,* Vol. 54, No. 2, 88-114, doi: 10.1016/S0007-8506(07)60021-3.

[11] Hamilton, K., Jeswiet, J. (2010). Single point incremental forming at high feed rates and rotational speeds: Surface and structural consequences, *CIRP Annals,* Vol. 59, No. 1, 311-314, doi: 10.1016/j.cirp.2010.03.016.

[12] Oleksik, V., Pascu, A., Deac, C., Fleacă, R., Bologa, O., Racz, G. (2010). Experimental study on the surface quality of the medical implants obtained by single point incremental forming, *International Journal of Material Forming*, Vol. 3, Supplement 1, 935-938, doi: 10.1007/s12289-010-0922-x.

[13] Bhattacharya, A., Maneesh, K., Venkata Reddy, N., Cao, J. (2011). Formability and surface finish studies in single point incremental forming, *Journal of Manufacturing Science and Engineering*, Vol. 133, No. 6, doi: 10.1115/1.4005458.

[14] Palumbo, G., Brandizzi, M. (2012). Experimental investigations on the single point incremental forming of a titanium alloy component combining static heating with high tool rotation speed, *Materials & Design*, Vol. 40, 43-51, doi: 10.1016/j.matdes.2012.03.031.

[15] Ambrogio, G., Filice, L., Gagliardi, F. (2012). Improving industrial suitability of incremental sheet forming process, *The International Journal of Advanced Manufacturing Technology*, Vol. 58, No. 9-12, 941-947, doi: 10.1007/s00170-011-3448-6.

[16] Silva, P.J., Leodido, L.M., Silva, C.R.M. (2013). Analysis of incremental sheet forming parameters and tools aimed at rapid prototyping, *Key Engineering Materials,* Vol. 554-557, 2285-2292, doi: 10.4028/www.scientific.net/KEM.554-557.2285.

[17] Lasunon, O.U. (2013). Surface roughness in incremental sheet metal forming of AA5052, *Advanced Materials Research,* Vol. 753-755, 203-206, doi: 10.4028/www.scientific.net/AMR.753-755.203.

[18] Filice, L., Fratini, L., Micari, F. (2002). Analysis of material formability in incremental forming, *CIRP Annals*, Vol. 51, No. 1, 199-202, doi: 10.1016/S0007-8506(07)61499-1.

[19] Skjoedt, M., Hancock, M.H., Bay, N. (2007). Creating helical tool paths for single point incremental forming, *Key Engineering Materials,* Vol. 344, 583-590, doi: 10.4028/www.scientific.net/KEM.344.583.

[20] Lu, B., Chen, J., Ou, H., Cao, J. (2013). Feature-based tool path generation approach for incremental sheet forming process, *Journal of Materials Processing Technology*, Vol. 213, No. 7, 1221-1233, doi: 10.1016/j.jmatprotec.2013.01.023.

[21] Liu, Z., Liu, S., Li, Y., Meehan, P.A. (2014). Modeling and optimization of surface roughness in incremental sheet forming using a multi-objective function, *Materials and Manufacturing Processes*, Vol. 29, No. 7, 808-818, doi: 10.1080/10426914.2013.864405.

[22] Mugendiran, V., Gnanavelbabu, A., Ramadoss, R. (2014). Parameter optimization for surface roughness and wall thickness on AA5052 aluminium alloy by incremental forming using response surface methodology, *Procedia Engineering*, Vol. 97, 1991-2000, doi: 10.1016/j.proeng.2014.12.442.
[23] Lu, B., Fang, Y., Xu, D.K., Chen, J., Ou, H., Moser, N.H., Cao, J. (2014). Mechanism investigation of friction-related effects in single point incremental forming using a developed oblique roller-ball tool, *International Journal of Machine Tools and Manufacture*, Vol. 85, 14-29, doi: 10.1016/j.ijmachtools.2014.04.007.
[24] Azevedo, N.G., Farias, J.S., Bastos, R.P., Teixeira, P., Davim, J.P., de Sousa, R.J.A. (2015). Lubrication aspects during single point incremental forming for steel and aluminum materials, *International Journal of Precision Engineering and Manufacturing*, Vol. 16, No. 3, 589-595, doi: 10.1007/s12541-015-0079-0.
[25] Reddy, N.V., Cao, J. (2008). Incremental sheet metal forming: A review, In: *Proceedings of the Indo-US workshop on Smart Machine Tools, Intelligent Manufacturing Systems at Multiscale Manufacturing*, PSG college of Technology, Combotore, India.
[26] Petek, A., Kuzman, K., Kopač, J. (2009). Deformations and forces analysis of single point incremental sheet metal forming, *Archives of Materials Science and Engineering*, Vol. 35, No. 2, 35-42.
[27] Buffa, G., Campanella, D., Fratini, L. (2013). On the improvement of material formability in SPIF operation through tool stirring action, *The International Journal of Advanced Manufacturing Technology*, Vol. 66, No. 9-12, 1343-1351, doi: 10.1007/s00170-012-4412-9.

APEM journal
Advances in Production Engineering & Management
Volume 13 | Number 4 | December 2018 | pp 417–428
https://doi.org/10.14743/apem2018.4.300
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Visual and optometric issues with smart glasses in Industry 4.0 working environment

Vujica Herzog, N.[a,*], Buchmeister, B.[a], Beharic, A.[b], Gajsek, B.[c]

[a]University of Maribor, Faculty of Mechanical Engineering, Maribor, Slovenia
[b]Healthcare Center dr. Adolfa Drolca Maribor, Maribor, Slovenia
[c]University of Maribor, Faculty of Logistics, Maribor, Slovenia

ABSTRACT

Smart glasses are a kind of Head Mounted Display (HMD) with great potential in Industry 4.0 working environments, where shop floor workers must be supplied with critical information in a timely, accessible and safe manner to be as productive as possible. Smart glasses collect data from a wireless network and project it on a tiny screen before the user's eye. Despite several benefits, such as hands-free access to computer-generated info, routeing to storage locations, eliminating the need to carry handheld scanners or written documents, there are also possible problems evidenced from the literature. HMD can cause headaches, pressure in the eyes, problems with focusing and difficulties with text reading. To study the addressed problems, a research was performed together with Ophthalmologists from Maribor Healthcare Centre. The effects of using Vuzix M300 Smart glasses on users' comfort during order picking activities were researched in a testing warehouse environment at the Faculty of Mechanical Engineering, Maribor. The testing period lasts four hours. Several ophthalmologic tests (visual acuity, contrast sensitivity, visual field testing and colour test) were performed before and after use of smart glasses. Results show that there are some statistically significant differences before and after use of smart glasses in users' visual acuity and, surprisingly, a high percentage of scotomas in the right eye (where the projection of smart glasses was performed) after use of smart glasses that cannot be overlooked.

ARTICLE INFO

Keywords:
Head-mounted display (HMD);
Smart glasses;
Industry 4.0;
Warehouse;
Manual order picking system

**Corresponding author:*
natasa.vujica@um.si
(Vujica Herzog, N.)

Article history:
Received 25 May 2018
Revised 19 November 2018
Accepted 23 November 2018

1. Introduction

Radical transformation of the manufacturing systems under the aegis of Industry 4.0 is facilitated by the concurrent development of disruptive technologies and the digital era. Factories are becoming smarter and more information-rich. For the first time, we have (evolving) technologies that can supply shop floor workers with critical information in a timely, accessible and safe manner to be as productive as possible. Smart glasses are a representative piece of equipment that makes that possible [1]. It lasted less than four years from the first public presentation of Google Glass to the first commercial use of smart glasses in the industrial environment. In the paper, we focus on manual order picking, as an example of a production working environment in a transition to Industry 4.0, and on the usability of smart glasses, as an example of industry 4.0 enabling technology.

Most order-picking systems used in practice are manual "picker to part" systems, and more than 80 % of all orders processed by warehouses are picked manually [2, 3]. The order picking

process, a process in which humans are routed by picking lists to items' storage locations to retrieve items for customers, is the most laborious and the most costly activity in a typical warehouse. With up to 55 % of the warehouse total operating costs [4], it is obvious why many companies are improving their order-picking tasks by using more efficient methods [5-7] and technologies [8, 9]. Since walking presents up to 50 % of the total picking time [4, 10], the logical way of improving this is to reduce or eliminate the unproductive walking, set-up, and searching time.

To improve efficiency in order picking operations, companies are experimenting increasingly with smart glasses [11], a kind of Head-Mounted Display (HMD). Powered by their own processor and battery, they collect data from a wireless network and project it onto a tiny screen incorporated into the glasses. From a user's perspective, the display looks like a full-sized display of text or graphic, overlaid on top of the "real world" scene viewed at the time. Workers benefit from hands-free access to computer-generated info, routeing to storage locations, eliminating the need to carry handheld scanners or written documents, thereby working conditions and productivity can be increased in parallel.

Despite the first pilot projects, technology is still developing. Theory [12] and practice [11] propose experiments with different products and applications before wider use in the warehouse. Some general recommendations presented in [13] can also be taken into account. A fundamental question is whether it could be harmful for the human eye to work a full day with smart glasses [14, 15].

2. Literature review

Systems using Head-Mounted Displays (HMDs) to support the order picking process are, in theory and practice, named Pick-by-vision systems. These systems are further divided according to (1) The ability to track user movements, and (2) The way of displaying information to a user in two subgroups [16]:

- Pick-by-vision (2D) systems (user position is not tracked, textual information in the form of a list of items or images is projected on the user's HMD);
- Pick-by-vision (AR) systems (use tracking, and make explicit use of Augmented Reality (AR) in a way that virtual objects are overlaid on the real-world environment and consist of the following parts [17]: Display, computer, input device and tracking system).

Smart glasses, also named as data glasses, are an example of HMD. In this paper, we focus on the possible harmful effects of the use of smart glasses.

Peli [18] researched Visual issues in the use of HMD already in 1990. Findings that base on a max of 20 minutes of tests in a laboratory environment, did not reveal any potential harmful effects, except not recommending use while driving. Six years later, Peli [12] wrote that the concerns about possible harmful effects are accompanying the introduction of almost any new wide-use-technology and HMD are not an exception. He concluded that it appears to be most appropriate to test each system separately. This will enable the developer to determine for each design that comfortable and safe use, by the target population and the intended use, is achievable.

Two years later, Peli [19] performed test sessions and measured the following visual parameters: (1) Accommodative status by refraction (auto refractor); (2) Binocular (OU) visual acuity at distance (6 m) with habitual correction; (3) Fixation disparity (lateral and vertical) at distance; (4) Stereoacuity at near distance (40 cm); (5) Phoria (lateral and vertical) at distance, and near (cover-test with prism-bar and Von-Graefe in the phoropter); (6) Vergence (horizontal and vertical) at distance and near; (7) Accommodative reserve by Fuse Cross Cylinder (FCC); (8) Convergence reserve, measured by negative and positive relative accommodation (NRA and PRA, respectively); (9) TBU time; (10) Contrast sensitivity at distance (OU) at three spatial frequencies (2, 3, and 6 c/deg). The reported data show no harmful or statistically significant changes to the visual system associated with use of the i-glasses HMD, in either *stereo* or *mono* mode, relative to the use of a desktop CRT display.

Researchers often use NASA Task Load Index (NASA-TLX), a subjective, multidimensional assessment tool that rates perceived workload in order to assess pick-by-vision systems or other

aspects of performance. It has been cited in over 4,400 studies, highlighting the influence the NASA-TLX has had in human factors' research [20]. Regarding user strain, Schwerdtfeger *et al.* [16] found that even though they have uncomfortable HMD headbands, a backpack to carry, and non-addressable display focal planes, their system did not cause a higher general user strain than the conventional paper list. Nevertheless, the discomfort questionnaire shows that improvements of the display devices are necessary to reduce the potential for headaches.

After 2010, smart glasses have been adopted as a safe enough technology for use in pilot projects. Since then, we have begun to encounter more research that explores different technical designs and combinations of different technologies to achieve optimum work results. Ergonomic aspects are still in the background of productivity studies.

In 2014, the number of publications on smart glasses and HMD topics started to increase markedly, according to the World of Science (WoS), from 11 in 2014 to 46 in 2018. From 168 published papers, 44.6 % are from the Computer Science research area, 39.8 % from Engineering and 16.07 % from Ophthalmology.

With the rapid development of mobile Head-Mounted Display (HMD), the problem of visual discomfort and visual fatigue caused by watching Virtual Reality (VR) contents became a crucial concern for consumers and manufacturers, especially given that the casing of a mobile HMD keeps the phone at a specified distance from the lenses that is close to the eyes [21]. In this regard, Jungmin *et al.* [21] conducted both subjective and objective measures to evaluate visual discomfort and visual fatigue caused by watching HMD and smartphones. Participants answered a Simulator Sickness Questionnaire (SSQ) and went through optometric tests that measure tear break-up time, spherical equivalent, and contrast sensitivity. Experimental results show that HMD causes more eye dryness compared to smartphones.

Klein-Theyer *et al.* [14] agree that the implementation of near-eye display devices is promising for the future of order picking systems and in various other workplace scenarios. However, in 2017, the workload associated with the use of a visually guided commissioning system had not yet been investigated. Authors investigate ocular comfort, ocular surface and tear function parameters before and after the completion of a task using either a visual- or a voice-guided picking solution. Recent publications indicate that up to 90 % of computer users experience ocular discomfort after prolonged computer use, and approximately 10 % of visual display unit users have severe complaints [22-25]. The visual analogue scale values were increased significantly with the visual system when compared to the voice system, and with the visual system after the work session had finished (i.e., pre- vs post-task), which suggests that visually guided picking solutions may influence ocular comfort adversely. The analysis of the objective data revealed a significant decrease in the tear break-up time values of the right eyes, and a minor decrease in the values of the left eyes, following the completion of the visually guided picking task. The tear break-up time values for the voice-guided condition remained stable (right eyes) or even increased (left eyes) after the work.

From a chronological review of scientific articles, we can conclude that researchers have only begun to study the impact of using HMD, or more precisely smart glasses, on people's vision. Researchers agree that smart glasses have a potential to be used more widely in manual order picking systems. Ophthalmologic studies are rare, and based mostly on short-term use of HMDs, less than an hour. Although most authors evaluate a pick-by-vision system with HMDs as competitive, productive and promising technology, which hold large potential in the future, questions linked to the effects of long-term use are still unanswered [15].

3. Materials and methods

We tested the effects of using Vuzix M300 Smart glasses on users' comfort during order picking activities in a testing warehouse environment at the Faculty of Mechanical Engineering, Maribor. The protocol of performed research is described below, and summarised in Fig. 1. 14 persons, mostly students, tested selected Head-Mounted Displays (HMD), owned by the company Špica International. The testing period lasted four hours. Before and after use of smart glasses several ophthalmologic tests (visual acuity, contrast sensitivity, visual field testing and colour test) were performed, therefore we got 28 measurements altogether.

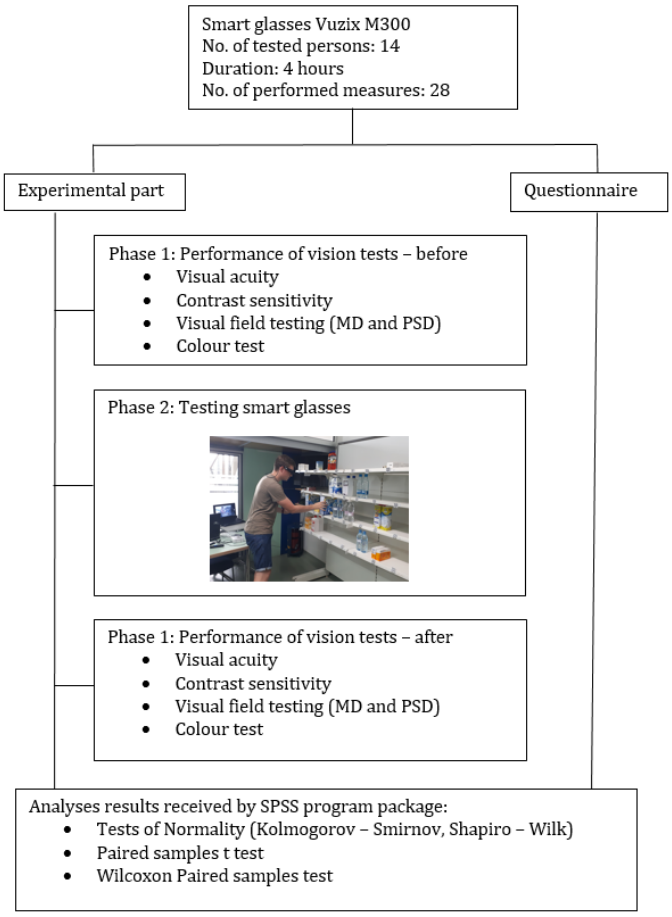


Fig. 1 Experimental protocol – Testing Vuzix M300 smart glasses

3.1 Performed ophthalmologic tests

Visual acuity

Measurement of visual acuity is a sensitive test of the integrity of the visual system. It fulfils all standard criteria of a good screening test: Minimal cost or risk to the patient, measurement can be performed quickly and easily with little or no examiner training, there is a high prevalence of detectable abnormalities, and abnormalities are most often amenable to treatment. The goal in testing central visual acuity is to determine the best possible visual acuity in each eye. In most instances, either a standard printed Snellen eye chart is used, or a reading card such as, e.g. ETDRS chart (Early Treatment Diabetic Study). One eye at a time is tested with the fellow eye occluded.

In our study, visual acuity was measured with use of the Snellen table. It is less accurate and repeatable as an ETDRS chart. Thus, we decided to include the candidates who had the best corrected visual acuity of 0.7 or more. The best corrected visual acuity in emmetropic patients is 1.0. The reason for including a factor of 0.7 of visual acuity is also for better cooperation in the test of visual field (perimetry), and for excluding possible pathology which could impair the visual acuity. Visual acuity was measured separately for right and left eyes.

Pelli Robson contrast sensitivity

Contrast sensitivity was performed by using the Pelli-Robson table (Fig. 2) which is fast and reliable enough for our study. In the literature it is used as a factor connected with loss of Retinal Nerve Fibre Layer (RNFL), and, thus, with impairment of visual functions. We tested each eye separately and also binocularly. During the testing, we encouraged the tested subject to concentrate himself to say as much as possible. The normal result is a log value of 1.95 or 2.0. Values less than 1.8 could indicate improper contrast vision in bad visual conditions.

Fig. 2 Pelli Robson table

Visual field testing

A visual field test is an eye examination that can detect dysfunction in central and peripheral vision, which may be caused by various medical conditions. Visual field testing can be performed clinically by keeping the subject's gaze fixed while presenting objects at various places within their visual field.

Visual field testing of 30 degrees of the central visual field was performed with a computer static perimetry full threshold algorithm using an OCTOPUS machine (Fig. 3) at standardised illumination parameters. 30 degree visual field testing could show possible scotomas (visual field defects) in the areas where the glasses were projected. For reliability of perimetry testing, the cooperation of the tested individual is very important. At a loss of fixation during the testing of more than 20 %, and false positive or negative below 15 %, the test result is inaccurate, and not reliable enough to say with 95 % possibility that scotoma actually exists. Together with visual acuity better than 0.7 and ability of tested individuals to be concentrated during the examination, makes the result more reliable regarding the new scotomas after the use of glasses.

Ishihara colour test

The colour vision testing was performed with use of Ishihara tables, which are used widely in clinical examination, and can indicate the presence of colour vision defects reliably, especially for results with more than 10 % of failures. The advantages using Ishihara tables is their ease of use and speed of performing, so the tested individual is not bored and losing his concentration. In the literature, the Hue 100 test is also performed mainly to quantify and to distinguish what kind of anomaly (protanomaly, deuteranomaly, and tritanomaly) is present in the tested individual. In our study, the colour vision was important, mainly to compare the results before and after the use of glasses.

Fig. 3 Octopus perimeter (Source: mandarinoptomedic.com)

3.2 Statistical analyses

All statistical analyses were performed through the Statistical Package for Social Sciences (SPSS) version 25. Within the descriptive statistics, the mean, Standard Deviations and standard error mean were calculated for all measurements.

The population normality was verified by using Kolmogorov-Smirnov and Shapiro-Wilk tests. Differences between performed tests before and after have been verified by the *t*-test for paired samples, or the Wilcoxon signed rank test, depending on the distribution of changes. Since the

data were not normally distributed in certain cases, we used the paired samples *t*-test for the normal and Wilcoxon test for the non-normal data within the group comparison. Additionally, effect sizes were calculated, to have a standardised measure of the size of the effect we observed, and to be able to compare results to other studies if they appeared.

In order to compare the results gathered with all performed ophthalmologic tests, the null hypothesis was stated that no changes are expected before and after use of smart glasses: $H_o : \mu_{\text{test before}} = \mu_{\text{test after}}$

The values were considered statistically significant at $p \leq 0.05$.

4. Results and discussion

4.1 Visual acuity

Normal visual acuity for a healthy human is 1.0. It means that humans see clear optotypes with standard values at a defined distance (6 m). In our study (Table 1), the majority of included candidates had visual acuity more than 0.7 that we set as a limit value. Only three of the tested persons had visual acuity less than 0.7.

Table 1 Visual acuity before and after using smart glasses

No. of testee	*VARb*	*VARa*	*VALb*	*VALa*
1	0.8	0.8	0.8	0.8
2	1.0	1.0	1.0	1.0
3	1.0	1.0	1.0	1.0
4	0.25	0.25	0.4	0.4
5	0.63	0.5	0.4	0.32
6	1.0	0.8	1.0	0.8
7	1.0	1.0	1.0	1.0
8	1.0	0.8	1.0	1.0
9	0.8	0.63	0.8	0.63
10	1.0	1.0	1.0	1.0
11	1.0	1.0	1.0	1.0
12	0.4	0.4	0.8	0.63
13	0.8	0.8	0.8	0.8
14	0.8	0.63	0.8	0.63

Note: VARb – Visual Acuity, right eye, before; *VARa* – Visual Acuity, right eye, after; *VALb* – Visual Acuity, left eye, before; *VALa* – Visual Acuity, left eye, after

4.2 Pelli Robson contrast sensitivity

The lowest contrast at which a tested person can read three letters of the same group on a Pelli Robson table determines the logarithm evaluation of contrast sensitivity (Table 2).

Table 2 Contrast sensitivity before and after using smart glasses

No. of testee	*CSRb*	*CSRa*	*CSLb*	*CSLa*	*CSbothb*	*CSbotha*
1	1.65	1.65	1.65	1.65	1.80	1.80
2	1.65	1.65	1.65	1.65	1.80	1.80
3	1.65	1.65	1.65	1.65	1.80	1.80
4	1.50	1.50	1.50	1.50	1.65	1.65
5	1.65	1.65	1.50	1.50	1.65	1.65
6	1.65	1.65	1.65	1.50	1.65	1.65
7	1.65	1.65	1.65	1.65	1.80	1.80
8	1.65	1.65	1.65	1.65	1.80	1.80
9	1.65	1.35	1.65	1.35	1.80	1.65
10	1.65	1.65	1.65	1.50	1.95	1.95
11	1.65	1.65	1.65	1.65	1.80	1.80
12	1.65	1.50	1.65	1.50	1.80	1.50
13	1.65	1.65	1.65	1.65	1.80	1.80
14	1.50	1.50	1.50	1.50	1.80	1.80

Note: CSRb – Contrast sensitivity, right eye, before; *CSRa* – Contrast sensitivity, right eye, after; *CSLb* – Contrast sensitivity, left eye, before; *CSLa* – Contrast sensitivity, left eye, after; *CSbothb* – Contrast sensitivity, both eyes, before; *CSbotha* – Contrast sensitivity, both eyes, after

Value 2.0 presents normal contrast sensitivity, or 100 %. Values less than 1.5 show visual handicap, and values less than 1.0 greater visual impairment. Results in Table 2 show that all tested candidates have contrast sensitivity equal to or greater than 1.5, which means that they can be treated as persons with normal contrast sensitivity.

4.3 Visual field testing

For visual field testing we used Octopus perimeter. First measurements were performed with the programme Treshold 30-2. The measurement was performed only on the right eye (this is the eye where the smart glasses display was). To examine changes in the visual field two parameters were used: *MD* (Mean Deviation) and *PSD* (Pattern Standard Deviation), which is used for quantifying the differences in sensitivity in the visual field. The purpose of our study here was to see if the difference of *MD* and *PSD* was significant.

Mean deviation

The average of deviations across all test locations is referred to as the Mean Deviation (*MD*). Subjects who are able to see dimmer stimuli than others of similar age and race, will have positive values for their *MD*, while subjects who require brighter stimuli will have negative *MD* values [26]. *MD* values for reliable tests typically range from +2 dB to –30 dB.

Results in the Table 3 show that the mean value of Mean Deviation before work with smart glasses was –2.54, and after 4-hours of using smart glasses –2.67. It means that the value after 4-hours working with smart glasses was lower by 5 % (Δ = 0.13).

Pattern standard deviation

Visual field loss in glaucoma is frequently non-uniform, and, thus, a measure which quantifies irregularities is desirable [26]. Pattern Standard Deviation (*PSD*) measures irregularity by summing the absolute value of the difference between the threshold value for each point and the average visual field sensitivity at each point (equal to the normal value for each point + the *MD*). Visual fields with the age-normal sensitivity at each point will have a *PSD* of zero, as will visual fields in which each point is depressed uniformly from the age-normal value. Thus, the largest *PSD* will be registered for focal, deep visual field defects. Near normal and severely damaged visual fields will both have low *PSD* [26]. Results of our tests are presented in Table 3.

Table 3 *MD* and *PSD* before and after using smart glasses

No. of testee	*MDb*	*MDa*	*PSDb*	*PSDa*
1	-2.25	-2.16	3.71	2.47
2	-0.71	-1.05	2.01	2.99
3	-4.80	-3.09	3.24	2.54
4	-5.21	-3.55	2.42	2.32
5	0.29	-2.87	3.10	3.19
6	-5.27	-4.48	3.45	3.02
7	-1.68	-2.77	2.37	2.33
8	-1.45	-3.99	1.97	2.18
9	-0.69	-2.63	1.79	2.61
10	-1.39	-1.07	2.09	2.37
11	-3.32	-3.79	2.15	3.15
12	-1.48	-2.29	2.34	1.99
13	-1.99	-3.15	2.09	3.21
14	-5.64	-0.49	3.62	2.34

Note: MDb – Mean Deviation, before; *MDa* – Mean Deviation, after; *PSDb* – Pattern Standard Deviation, before; *PSDa* – Pattern Standard Deviation, after

4.4 Ishihara colour test

We used 15 colour plates of the Ishihara test (numbers, including the test plates) with 12 or more correct indicating normal colour vision. Again, we wanted to measure the difference in colour vision perception before and after the use of smart glasses. In our result, there was some discrepancy regarding the individual possible mild colour anomaly, and regarding the individual

concentration during the test. Thus, the majority of tested persons had normal colour vision, except two individuals where an impairment in colour vision was indicated (Table 4).

We also studied the possible scotomas in the area of the visual field where the smart glasses had a projection. In some cases, scotomas were present after the use of smart glasses. That might indicate that glasses can cause some vision impairment after use, projecting in the same quadrant, lowering the sensitivity in that area of the visual field (Table 4).

The Treshold 30-2 programme was performed on the right eye, which is the eye where the display was. If we found a scotoma, we marked it as 1, and if the scotoma was not present, we marked it as 0. Results of our test show that before using smart glasses no tested persons had scotomas, but after using smart glasses 6 persons had it. It presents 43 % of those tested, therefore, we can justify that the presence of scotoma in the right eye can be the result of load caused by using smart glasses. To verify the results gained with the Treshold 30-2 programme we also performed the Driver's licence test for both eyes. According to the Driver's license test, we found two scotomas after using smart glasses, which is less than with the Treshold 30-2 programme. This could also be caused by the complexity of the performed research and the duration. Treshold 30-2 lasts 15 minutes and the Driver's license test only 4 minutes. In both tests the concentration of the tested persons is very important, and can influence the results.

Table 4 Colour test and visual field test before and after using smart glasses

			Treshold 30-2		Driver's licence	
No. of testee	*CTb*	*CTa*	*SCOb*	*SCOa*	*SCOb*	*SCOa*
1	15	15	0	0	0	0
2	3	3	0	0	0	0
3	15	15	0	0	0	0
4	15	15	0	1	0	0
5	15	15	0	0	0	0
6	1	1	0	1	0	1
7	15	15	0	0	0	0
8	15	15	0	1	0	0
9	15	15	0	0	0	0
10	15	15	0	1	0	0
11	15	15	0	1	0	0
12	15	15	0	0	0	0
13	15	15	0	1	0	1
14	15	15	0	0	0	0

Note: CTb – Colour test, before; *CTa* – Colour test, after; *SCOb* – Scotoma, before; *SCOa* – Scotoma, after

4.5 Statistical analysis results

Results of statistical analysis for all the performed analyses are summarised in Table 5. Statistically significant differences with $p \leq 0.05$ are in bold.

Table 5 Results of descriptive statistics, *t*-test for paired samples and Wilcoxon test

	Before			After				Wilcoxon	Effect
	Mean	*SD*	St. Err. Mean	Mean	*SD*	St. Err. Mean	*t*-test *p*	test *p*	size
VAR	0.820	0.242	0.064	0.757	0.245	0.065	**0.020**	**0.041**	**-0.38**
VAL	0.842	0.210	0.056	0.786	0.234	0.062	**0.024**	**0.039**	**-0.39**
CSR	1.628	0.054	0.014	1.596	0.095	0.025	0.189	0.18	-
CSL	1.617	0.063	0.017	1.564	0.096	0.025	0.055	0.059	-
CS	1.778	0.080	0.021	1.746	0.111	0.029	0.189	0.18	-
CT	13.143	4.737	1.266	13.143	4.737	1.266	-	1.00	-
MD	-2.542	1.950	0.521	-2.670	1.171	0.313	0.822	0.470	-
PSD	2.596	0.675	0.180	2.622	0.409	0.109	0.904	0.975	-
ST	0	0	0	0.428	0.513	0.137	**0.008**	**0.014**	**0.45**

Note: VAR – Visual Acuity, right eye; *VAL* – Visual Acuity, left eye; *CSR* – Contrast sensitivity, right eye; *CSL* – Contrast sensitivity, left eye; *CS* – Contrast sensitivity of both eyes; *CT* – Colour test; *MD* – Mean Deviation; *PSD* – Pattern Standard Deviation; *ST* – Scotoma Threshold 30-2

Results of visual acuity for the right eye show that the mean value before work with smart glasses was 0.82, and after 0.76. Visual acuity was lower after using smart glasses for Δ = 0.062 (7.6 %). It means that tested persons' sight was weaker after use of smart glasses than before. Since the results of both tests (*t* test and Wilcoxon test) show significant difference, the null hypothesis $H_o : \mu_{VARbefore} = \mu_{VARafter}$ should be disproved. The effect size is $r = -0.38$. This represents a medium to large change in visual acuity for the right eye.

For the left eye, the results of visual acuity are similar as for the right eye, but there are slight differences. The mean value before work with smart glasses was 0.84, and after 0.76. Visual acuity after using smart glasses was also lower by Δ = 0.056 (6.7 %), meaning that tested persons' sight was weaker after use of smart glasses. According to the results of the *t* and Wilcoxon tests, the null hypothesis $H_o : \mu_{VARbefore} = \mu_{VARafter}$ should be disproved. The effect size is $r = -0.39$. This represents a medium to large change in visual acuity for the left eye.

Comparison between right and left eye show that visual acuity of the right eye reduced by 7.6 %, and 6.7 % for the left eye. All tested persons had the visual display of the smart glasses in front of their right eye, therefore the weaker sight on the right eye could be caused by using smart glasses.

The mean value of contrast sensitivity for the right eye before work with smart glasses was 1.63, and after 1.60. Contrast sensitivity was lower after using smart glasses by Δ = 0.03 (2 %). It means that the tested persons' contrast sensitivity was weaker after use of smart glasses than before. But the results of both tests (*t* and Wilcoxon test) do not show significant difference; *p* values are greater than 0.05, therefore the null hypothesis $H_o : \mu_{VARbefore} = \mu_{VARafter}$ cannot be disproved.

A similar situation is evidenced for the left eye, where the mean value of contrast sensitivity changed from 1.62 to 1.56, Δ = 0.05 (3 %), meaning that the tested persons' contrast sensitivity was weaker after the use of smart glasses than before. Results of both tests (t and Wilcoxon test) show no significant difference, therefore, the null hypothesis $H_o : \mu_{VARbefore} = \mu_{VARafter}$ cannot be disproved.

The mean value of contrast sensitivity for both eyes changed from 1.78 to 1.75, which means that contrast sensitivity was lower after using smart glasses by Δ = 0.03 (1.8 %). Tested persons perceived lower contrast sensitivity after using smart glasses, but, since the measured value was not lower than 1.5, use of smart glasses is not harmful for eye contrast sensitivity [27]. The mean value of perceiving contrast sensitivity was binocularly better than monocular, which is also verified with other studies [28].

Results of both tests (*t* and Wilcoxon test) gave us *p* values greater than 0.05, therefore, the null hypothesis $H_o : \mu_{VARbefore} = \mu_{VARafter}$ cannot be disproved.

The mean value of the Mean Deviation test (*MD* test) of visual field is -2.54 before using smart glasses and -2.67 after. The value after 4-hours use of smart glasses was lower by Δ = 0.13 (5 %). Regarding the values of the *t* and Wilcoxon tests, the null hypothesis $H_o : \mu_{VARbefore} = \mu_{VARafter}$ cannot be disproved.

Our results were lower than the normal range before testing, which can be influenced by the patient's psychological state, concentration and cooperation during the test. We must be aware of the state when the tested individual performs the first test of the visual field. Besides that, our results statistically did not show a significant difference before and after the test.

The mean value of the Pattern Standard Deviation test (*PSD* test) before work with smart glasses was 2.6, and after 2.62. The tested value after 4-hours use of smart glasses was slightly higher by Δ = 0.03, or less than 1 %. Regarding the values of the *t* and Wilcoxon tests, the null hypothesis $H_o : \mu_{VARbefore} = \mu_{VARafter}$ cannot be disproved.

High values of PSD represent the scotomas [26]. The Treshold 30-2 Programme confirmed the increased level of scotoma in the right eye, but the Driver's license test gave us a lower level of scotomas. Even the second test couldn't confirm the results of Treshold 30-2 totally, so it can be concluded that the presence of scotoma in the right eye can be the result of load caused by using smart glasses. The results of the *t* and Wilcoxon tests had *p* values less than 0.05, therefore, the null hypothesis $H_o : \mu_{VARbefore} = \mu_{VARafter}$ should be disproved. The effect size is $r = 0.45$. This represents a medium to large change in the level of scotoma.

The results of the performed ophthalmologic tests (visual acuity, contrast sensitivity, visual field testing and colour test), show that there are some statistically significant differences between tests results performed before and after use of smart glasses that cannot be overlooked. The difference between the results of visual acuity for the left and right eyes is small, but, in both cases, the visual acuity is lower after use of smart glasses. The contrast sensitivity and colour test did not show any statistically significant differences, but the additional test of the visual field did.

From the results of the visual field test we found out that none of the tested individuals had the scotoma (dysfunction in central and peripheral vision) in the right eye (inferior quadrant where the projection of smart glasses was performed) before using smart glasses. After the test, scotomas were present in the same quadrant in 43 % of cases. This might indicate that use of smart glasses for four hours during work can cause scotomas and, thus, impairment in the visual field and vision.

The percentage of scotomas in the right eye after using smart glasses is high, therefore, we tried to find out if there are any other influence parameters that should be considered. It is known that perimetry is a subjective testing, affected mainly by the psychological state of the individual and by their cooperation and concentration. Therefore, this might be the influential factor. We also did only one measure before and after per tested person, because the test is time consuming, and it is known from the literature [26] that perimetry gives the best results usually at a second testing, when the tested person is more familiar with the procedure. The size of the tested group was relatively small, but even though it was performed as a pilot test, results were somehow surprisingly high. We also did the additional Driver's Licence test, which is faster and, thus, less affected by the motivation of tested individual. In the Driver's licence test the presence of scotomas was lower, but they were still present in 14 % of cases.

5. Conclusion

With the intent to create added value, warehouses must be aligned with modern industry trends, as well as with novel business approaches enabled by modern forms of organisation and ICT [29-31].

Systems using Head-Mounted Displays (HMDs) are still developing, and there is a lack of research concerning human comfort during a full work day with this kind of equipment. In our research, we tried to answer the question whether use of smart glasses could be harmful for the human eye or not during four-hour use.

Based on the results of all the performed tests, we can conclude that use of smart glasses has an effect on the user's vision, and, therefore, further research would be of benefit before implementing it in warehouses as a part of everyday equipment for workers. During our research, a number of questions appeared that could be addressed in the future:

- Visual acuity during the workday; Does normal visual acuity of a healthy human change during the day?
- Appearance of scotoma; we researched the presence of scotomas only in the right eye and, therefore, we couldn't compare the results with results of the left eye.
- Appearance of scotoma during the day.
- Appearance of scotoma regarding the kind of work; additional research would be of benefit if certain kinds of work could be identified that have greater prevalence for scotoma formation.
- Identical research could be performed, but with the projection of smart glasses in the left inferior quadrant.
- Does the appearance of scotoma affect a driver's abilities?

Our pilot test was performed with limited resources on 14 persons. All tests were performed twice, before and after working with smart glasses, therefore, our conclusions based on 28 measures. In the future, rigorous, empirically based research, performed on a greater number of tested persons, could help to clear up doubts that still exist concerning the presented topic.

Acknowledgement

This work was supported by the Slovenian Research Agency in the framework of Grant P2-0190. The research was funded by The Public Scholarship, Development, Disability and Maintenance Fund of the Republic of Slovenia.

References

[1] Ballard, B. (2017). How shop floor smart glasses create a more connected workforce, from *https://www. manufacturing.net/article/2017/01/how-shop-floor-smart-glasses-create-more-connected-workforce*, accessed November 16, 2018.
[2] De Koster, R., Le-Duc, T., Roodbergen, K.J. (2007). Design and control of warehouse order picking: A literature review, *European Journal of Operational Research*, Vol. 182, No. 2, 481-501, doi: org/10.1016/j.ejor.2006. 07.009.
[3] Burinskiene, A. (2015). Optimising forklift activities in wide-aisle reference warehouse, *International Journal of Simulation Modelling*, Vol. 14, No. 4, 621-632, doi: 10.2507/IJSIMM14(4)5.312.
[4] Tompkins, J.A., White, J.A., Bozer, Y.A., Tanchoco, J.M.A. (2003). *Facilities planning,* Third edition, Hoboken, New Jersey, John Wiley, USA.
[5] Banduka, N., Veža, I., Bilić, B. (2016). An integrated lean approach to process failure mode and effect analysis (PFMEA): A case study from automotive industry, *Advances in Production Engineering & Management*, Vol. 11, No. 4, 355-365, doi: 10.14743/apem2016.4.233.
[6] Gholamian, M.R., Heydari, M. (2017). An inventory model with METRIC approach in location-routing-inventory problem, *Advances in Production Engineering & Management*, Vol. 12, No. 2, 115-126, doi: 10.14743/apem2017. 2.244.
[7] Berlec, T., Kleindienst, M., Rabitsch, C., Ramsauer, C. (2017). Methodology to facilitate successful lean implementation, *Strojniški Vestnik – Journal of Mechanical Engineering*, Vol. 63, No. 7-8, 457-465, doi: 10.5545/sv-jme. 2017.4302.
[8] Dukic, G., Rose, L., Gajsek, B., Opetuk, T., Cajner, H. (2018). Space time and ergonomic assessment of order-picking using vertical lift modules, In: *Proceedings of 14th International Conference on Industrial Logistics*, Ben-Gurion University, Beer-Sheva, Israel, 68-74.
[9] Karabegović, I., Karabegović, E., Mahmić, M., Husak, E. (2015). The application of service robots for logistics in manufacturing processes, *Advances in Production Engineering & Management*, Vol. 10, No. 4, 185-194, doi: 10.14743/apem2015.4.201.
[10] Dukić, G., Oluić, Č. (2004). Order-picking routing policies: Simple heuristics, advanced heuristics or optimal algorithm, *Strojniški Vestnik – Journal of Mechanical Engineering*, Vol. 50, No. 11, 530-535.
[11] Ames, B. (2017). Smart glasses get a second look from warehouses, from *http://www.dcvelocity.com/articles/ 20170213-smartglasses-get-a-second-look-from-warehouses/*, accessed October 23, 2018.
[12] Peli, E. (1996). Visual and optometric issues with head-mounted displays, In: *Proceedings of the IS&T Optics & Imaging in the Information Age*, The Society for Imaging Science and Technology, Rochester, New York, 364-369.
[13] Eraslan, E., Can, G.F., Atalay, K.D. (2016). Mental workload assessment using a fuzzy multi-criteria method, *Tehnički Vjesnik – Technical Gazette*, Vol. 23, No. 3, 667-674, doi: 10.17559/TV-20140401112509.
[14] Klein-Theyer, A., Horwath-Winter, J., Rabensteiner, D.F., Schwantzer, G., Wultsch, G., Aminfar, H., Heidinger, A., Boldin, I. (2016). The impact of visual guided order picking on ocular comfort, ocular surface and tear function, *PLoS One*, Vol. 11, No. 6, doi: 10.1371/journal.pone.0157564.
[15] Josefsson, P., Lingegård, S. (2017). *Potential of smart glasses in a spare parts distribution center*, Master's thesis, Chalmers University of Technology, Gothenburg, Sweden.
[16] Schwerdtfeger, B., Reif, R., Gunthner, W.A., Klinker, G., Hamacher, D., Schega, L., Bockelmann, I., Doil, F., Tulmler, J. (2009). Pick-by-vision: A first stress test, In: *Proceedings of IEEE International Symposium on Mixed and Augmented Reality*, Orlando, Florida, USA, 115-124, doi: 10.1109/ISMAR.2009.5336484.
[17] Rammelmeier, T., Galka, S., Günthner, W.A. (2011). Active prevention of picking errors by employing pick-by-vision, In: *Proceedings of 4th International Doctoral Students Workshop on Logistics*, Otto-von-Guericke Universtät Magdeburg, Magdeburg, Germany, 79-83.
[18] Peli, E. (1990). Visual issues in the use of a head-mounted monocular display, *Optical Engineering*, Vol. 29, No. 8, 883-892, doi: 10.1117/12.55674.
[19] Peli, E. (1998). The visual effects of a head-mounted display (HMD) are not distinguishable from those of a desk-top computer display, *Vision Research*, Vol. 38, No. 13, 2053-2066, doi: 10.1016/S0042-6989(97)00397-0.
[20] Schuff, D., Corral, K., Turetken, O. (2011). Comparing the understandability of alternative data warehouse schemas: An empirical study, *Decision Support Systems*, Vol. 52, No. 1, 9-20, doi: 10.1016/j.dss.2011.04.003.
[21] Han, J., Bae, S.H., Suk, H.-J. (2017). Comparison of visual discomfort and visual fatigue between head-mounted display and smartphone, In: *Proceedings of the IS&T International Symposium on Electronic Imaging, Human Vision and Electronic Imaging,* Society for Imaging Science and Technology, 212-217, doi: 10.2352/ ISSN.2470-1173.2017.14.HVEI-146.
[22] Mocci, F., Serra, A., Corrias, G.A. (2001). Psychological factors and visual fatigue in working with video display terminals, *Occupational & Environment Medicine*, Vol. 58, No. 4, 267-271, doi: 10.1136/oem.58.4.267.
[23] Cole, B.L., Maddocks, J.D., Sharpe, K. (1996). The Effect of VDUs on the eyes: A report of a 6-year epidemiological study, *Optometry and Vision Science*, Vol. 73, No. 8, 512-528, doi: 10.1097/00006324-199608000-00001.

[24] Blehm, C., Vishnu, S., Khattak, A., Mitra, S., Yee, R.W. (2005). Computer vision syndrome: A review, *Survey of Ophthalmology*, Vol. 50, No. 3, 253-262, doi: 10.1016/j.survophthal.2005.02.008.
[25] Rosenfield, M. (2011). Computer vision syndrome: A review of ocular causes and potential treatments, *Ophthalmic & Physiological Optics*, Vol. 31, No. 5, 502-515, doi: 10.1111/j.1475-1313.2011.00834.x.
[26] Ramulu, P., Salim, S. (2017). Standard automated perimetry, from *http://eyewiki.aao.org/Standard_Automated_Perimetry*, accessed August 21, 2018.
[27] Parede, T.R.R., Torricelli, A.A.M., Mukai, A., Netto, M.V., Bechara, S.J. (2013). Quality of vision in refractive and cataract surgery, indirect measurers: Review article, *Arquivos Brasileiros de Oftalmologia*, Vol. 76, No. 6, 386-390, doi: 10.1590/S0004-27492013000600016.
[28] Karatepe, A.S., Köse, S., Eğrilmez, S. (2017). Factors affecting contrast sensitivity in healthy individuals: A pilot study, *Turkish Journal of Ophthalmology,* Vol. 47, No. 2, 80-84, doi: 10.4274/tjo.93763.
[29] Tepeš, M., Krajnik, P., Kopač, J. (2015). Framework proposition and technical guidelines for manufacturers of custom made tools, machinery and special equipment, *Tehnički vjesnik – Technical Gazette,* Vol. 22, No. 3, 581-590, doi: 10.17559/TV-20140114133656.
[30] Veza, I., Mladineo, M., Gjeldum, N. (2016). Selection of the basic lean tools for development of croatian model of innovative smart enterprise, *Tehnički Vjesnik – Technical Gazette,* Vol. 23, No. 5, 1317-1324, doi: 10.17559/TV-20160202120909.
[31] Işik, M.F., Haboğlu, M.R., Yilmaz, C., Yilmaz, E.N. (2018). Design and implementation of real-time monitoring and control system for distributed robotic systems supported with IOS/Android application, *Tehnički Vjesnik – Technical Gazette,* Vol. 25, No. 2, 423-428, doi: 10.17559/TV-20160125160226.

APEM journal
Advances in Production Engineering & Management
Volume 13 | Number 4 | December 2018 | pp 429–441
https://doi.org/10.14743/apem2018.4.301
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Multi-objective production planning model for equipment manufacturing enterprises with multiple uncertainties in demand

Liu, Y.F.[a,b,*], **Zhang, Q.S.**[a]

[a]School of Management, Shenyang University of Technology, Shenyang, P.R. China
[b]School of Management, Bohai University, Jinzhou, P.R. China

ABSTRACT

A production planning model with multiple uncertainties was established in this paper. Customers' demands for quantity, quality, delivery time and price are different. For an ambiguous number of customers, the expectation of the degree of satisfaction was determined by a triangular fuzzy number method. A trapezoidal fuzzy number method was used for customer prices to determine the expectation of satisfaction of the delivery date. Fuzzy intervals and interval numbers were used to describe quality uncertainty and price uncertainty, respectively. A multi-objective planning model was established, which consists of four objectives, namely, meeting customers' needs, minimizing costs, minimizing delivery time and maximizing corporate profits. Then, the non-dominated sorting genetic algorithm (NSGA-II) was implemented to simulate and solve the problem of uncertain optimization. This modal resolved multiple uncertainties in customer demand during the process of production planning for the equipment manufacturing enterprises. The results of the running showed and generated a series of Pareto solutions, which are consistent with the results of a multi-objective planning solution. Manufacturers can obtain the best production plans according to the company's production objective priority rules. Finally, the adaptability and feasibility of the model were verified.

ARTICLE INFO

Keywords:
Production planning;
Multiple uncertainties;
Manufacturing enterprise;
Multi-objective model;
Non-dominated sorting genetic algorithm (NSGA-II)

**Corresponding author:*
liuyfmail1979@163.com
(Liu, Y.F.)

Article history:
Received 5 June 2018
Revised 26 August 2018
Accepted 7 September 2018

1. Introduction

Equipment manufacturing companies need to have flexible manufacturing capabilities because of the diversity, individuation, and differentiation of customer needs, and their production plans must respond quickly to fluctuations in demand. When operating near system limits, unexpected modifications to production plans can be very challenging. Scholars and experts have used mathematical models to research the management of uncertain needs. Rahdar *et al.* built a two-stage, tri-level optimization model under the uncertainty of demand and date of delivery [1]. Ho *et al.* built a model that addresses the problem of capacity allocation for multi-variety products when customer needs are uncertain [2]. Shi *et al*. established a mathematical model to solve the uncertainty of demand and returns and adopted the Lagrangian relaxation method [3]. Diabat *et al.* established a joint network positioning inventory model. The model solved the uncertainty of demand and lead-time using simulated annealing and direct search methods [4]. Feng and Nagi used a scenario analysis approach to deal with uncertain demand. They established an optimiza-

tion model to solve the problem of cost changes brought about by uncertain demand [5]. Erdogan and Denton established a multistage stochastic linear model. This model solves the problem of uncertain service dates and customer numbers on a specific date [6]. Cho and Tang studied three sales strategies to solve the problem of uncertainty in supply and demand [7]. Aghezzaf *et al.* established a stochastic model to solve the cyclical demand uncertainty of the finished product [8]. Chica *et al.* established a multi-objective optimization assembly line balancing model to address the variability and uncertainty of the demand for mixed products for industrial scenarios [9]. Chica *et al.* set up a multi-objective optimization model for assembly line balancing to address changes in demand for different products [10]. Xu *et al.* created multi-objective decision-making methods [11-12]. Tang *et al.* solved problems by establishing nondeterministic models [13-14]. Galal *et al.* built simulation models with uncertain demand quantities and delivery cycles to address agricultural supply issues [15]. Said *et al.* established a distribution operation model under fuzzy demand [16]. Liu and Zhang considered dual uncertainties and established a multi-objective planning model [17]. This paper extends research on this basis.

The paper has four sections. In Section 2, we describe the problem, conditional assumptions, and variable symbols. In Section 3, we focus on multiple uncertainties of demand and adopt a scenario analysis method to establish a multi-objective model for equipment manufacturing enterprises. In Section 4, we solve the model by using an NSGA-II algorithm. In Section 5, we conduct a case analysis and present a discussion.

2. Problem descriptions and premise assumptions

2.1 Problem description

Due to the impact of issues such as raw material supply times and finite order periods, most equipment manufacturing enterprises manage their production according to the customers' wishes. However, the processes involved in signing, manufacturing, and fulfilling orders may not be consistent. Therefore, in cases of order modification and emergency orders, or changes in customer demand for quantities, varieties, qualities, delivery times and prices, manufacturing processes are affected, and cause disruptions in production plans and the scheduling of equipment. Fig. 1 is a description of the multiple uncertainties in customer demand for an equipment manufacturing enterprise.

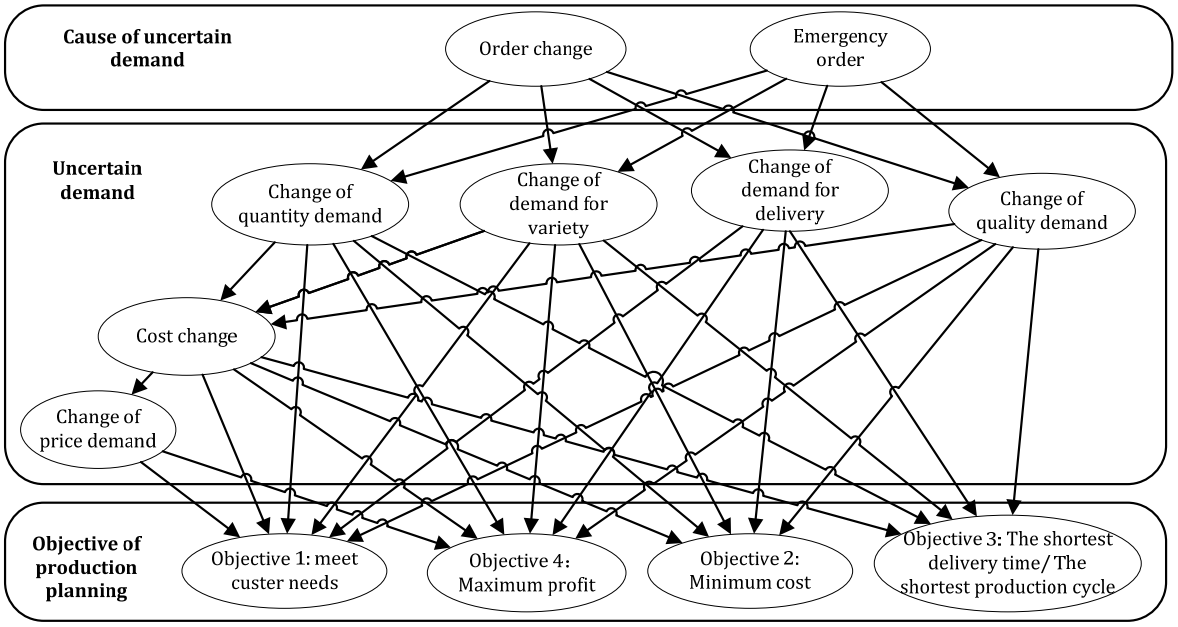


Fig. 1 Multiple uncertainties of customer demand

2.2 Premise assumptions

Based on the description of the multiple uncertainties of customer demands in Fig. 1, the following hypotheses are put forward:

1) Assume that the probability of order changes and emergency orders of the equipment manufacturing enterprises is known.
2) Assume that the manufacturing enterprise has production capacity constraints, inventory constraints, and resource acquisition capacity constraints.
3) Assume that order changes and urgent orders cause changes in the enterprises' production plans.

2.3 Variables and parameters description

The basic symbols and variables involved in the model and their meanings are in Table 1. The subscripts *i*, *j*, *t* in Table 1 are denoted as the *i*-th product, the *j*-th resource, and the *t*-th period, respectively. The variable *i* is an integer between 1 and *n*, and *j* is an integer between 1 and *m*.

Table 1 The meaning of basic symbols and variables

Symbols	Meanings
n	The number of product types of the enterprise.
m	The number of types of production resources.
N_{it}	The number of demands for original planned products.
N'_{ijt}	The number of resource requirements for the original plan.
R_{it}	The stock of the product.
R'_{jt}	Inventory of resource.
MR_i	The maximum stock of product.
MR'_j	The maximum stock of resource.
ln_{it}	The amount of stockout.
lt_{it}	The shortage time of production.
lq_{it}	The quality deviation.
lp_{it}	The price deviation.
C_{it}	The production cost.
RC_{it}	The production preparation costs.
MC_{it},	The minimum cost.
C'_{jt}	The acquisition cost.
P_{it}	The original planned price.
PM_{it}	The market price.
T_{it}	The total lead time to meet the demand.
RT_{it}	The production preparation lead-time.
ET_{it}	The lead time for emergency production.
TN_{it}	The customer's original demand time.
ls_i	The economic batch quantity.
Q_{it}	The original planned demand quality.
QF_{it}	The quality.
W_{it}	Ability to produce products
MW_{it}	Maximum ability to produce products.
W'_{jt}	Ability to produce or acquire resource.
MW'_{jt}	Maximum ability to produce or acquire resource.
E_{it}	The cost of opportunity loss.
F_{it}	The unit inventory holding cost.
G_{it}	The planned production quantity.
ΔN_{it}	Changes in the quantity of demand.
ΔTN_{it}	Changes in the time of demand.
ΔP_{it}	Changes in the price of demand.
ΔQ_{it}	Changes in the quality of demand
$\widehat{N}_{it}$	The quantity required of customer demand.
$\widehat{TN}_{it}$	The required time of the customer's demand.
$\widehat{P}_{it}$	The customer's demand price.
$\widehat{Q}_{it}$	The quality of the customer's demand.

3. Multi-objective production planning model

Taking into account the multiple uncertainties of the needs of equipment manufacturing companies' customers, this paper considers several objectives such as customer satisfaction, production cost, production cycle, and profit. We create a multi-objective production-planning model in Eq. 1.

$$\begin{gathered} min\, F(X\_vm) = \left(f_1(X\_vm), f_2(X\_vm), f_3(X\_vm), f_4(X\_vm)\right) \\ s.t. \quad X \in \Omega \\ G(X) = 0 \\ H(X) \le 0 \end{gathered} \tag{1}$$

In the model, f_1 is customer satisfaction, objective f_2 is production cost, objective f_3 is a production cycle objective, and f_4 is a corporate profit objective. X_vm is an optimization variable of the multi-objective model. It is a multidimensional vector matrix (SF, CS, PT, PF), where SF is the quantity vector matrix of the product produced by the enterprise, CS is the cost vector matrix of the product produced by the enterprise, PT is the production cycle vector of the products produced by the enterprise, PF is the profit vector matrix of the products produced by the enterprise, Ω is a feasible solution space, G is an equality constraint function, and H is an inequality constraint function.

3.1 Objective functions

In cases where a customer changes the order requirements and requires an emergency order, a scenario analysis method is used to describe the situation. The variables are: S_1 describes the scenario of the order change demand occurred, p_{s1} is the probability of occurrence of S_1, S_2 describes the scenario of the emergency order demand occurred, and p_{s2} is the probability of occurrence of S_2.

The S_1, S_2 combinations constitute S scenarios. S is a series of scenarios for equipment manufacturing enterprises' multi-variety production order change and emergency order demand. In addition, $p_s = \{p_{s1}, p_{s2}\}$, $p_{s1} + p_{s2} = 1$. According to the above description, the four production operation objectives are as follows:

Objective 1: High customer satisfaction is the objective, that is, low customer dissatisfaction is not an objective. The model is Eq. 2 in the form:

$$\begin{aligned} f_1 = min \sum_{s=1}^{S}\sum_{t=1}^{T} p_s \cdot \xi_{N1}^{ts-} + min \sum_{s=1}^{S}\sum_{t=1}^{T} p_s \cdot \xi_{T1}^{ts-} + min \sum_{s=1}^{S}\sum_{t=1}^{T} p_s \cdot \xi_{Q1}^{ts-} \\ + min \sum_{s=1}^{S}\sum_{t=1}^{T} p_s \cdot \xi_{P1}^{ts-} \end{aligned} \tag{2}$$

Subject to:

$$\sum_{i=1}^{n} (\widehat{N}_{it}^{s} - ln_{it}^{s}) / \sum_{i=1}^{n} N_{it}^{s} + \xi_{N1}^{ts-} - \xi_{N1}^{ts+} = 1, \forall t \in N, \forall s \in S \tag{3}$$

$$\sum_{i=1}^{n} (\widehat{TN}_{it}^{s} - lt_{it}^{s}) / \sum_{i=1}^{n} TN_{it}^{s} + \xi_{T1}^{ts-} - \xi_{T1}^{ts+} = 1, \forall t \in N, \forall s \in S \tag{4}$$

$$\sum_{i=1}^{n} (\widehat{P}_{it}^{s} - lp_{it}^{s}) / \sum_{i=1}^{n} P_{it}^{s} + \xi_{P1}^{ts-} - \xi_{P1}^{ts+} = 1, \forall t \in N, \forall s \in S \tag{5}$$

$$\sum_{i=1}^{n} (\widehat{Q}_{it}^{s} - lq_{it}^{s}) / \sum_{i=1}^{n} Q_{it}^{s} + \xi_{Q1}^{ts-} - \xi_{Q1}^{ts+} = 1, \forall t \in N, \forall s \in S \tag{6}$$

In Eqs. 2 to 6:

ξ_{N1}^{ts-} and ξ_{N1}^{ts+}, respectively, are ratios of the quantity of customer demand that has not been reached or exceeded;
ξ_{T1}^{ts-} and ξ_{T1}^{ts+}, respectively, are ratios of the delivery date of customer demand that has not been reached or exceeded.
ξ_{Q1}^{ts-} and ξ_{Q1}^{ts+}, respectively, are ratios of the quality of customer demand that has not been reached or exceeded.
ξ_{P1}^{ts-} and ξ_{P1}^{ts+}, respectively, are ratios of the price of customer demand that has not been reached or exceeded.

$\xi_{N1}^{ts-}, \xi_{N1}^{ts+}, \xi_{T1}^{ts-}, \xi_{T1}^{ts+}, \xi_{Q1}^{ts-}, \xi_{Q1}^{ts+}, \xi_{P1}^{ts-}, \xi_{P1}^{ts+}$ are all in the t-th period. In addition, the customer's preference for the quantity of demand, delivery time of demand, demand quality and demand price are different so the satisfaction coefficient of expectation is also different. Therefore, we can define one vector matrix as:

$$\beta = [\beta 1_k^{ts}, \beta 2_k^{ts}, \beta 3_k^{ts}, \beta 4_k^{ts}]$$

For any given $\forall t,\ s,\ k$ where $\beta 1_k^{ts}, \beta 2_k^{ts}, \beta 3_k^{ts}, \beta 4_k^{ts}$ the values are all [0,1], and

$$\beta 1_k^{ts} + \beta 2_k^{ts} + \beta 3_k^{ts} + \beta 4_k^{ts} = 1$$

$\beta 1_k^{ts}$ is the satisfaction expectation coefficients of the customer's demands for the quantity, $\beta 2_k^{ts}$ is the satisfaction expectation coefficients of the customer's demands for delivery time, $\beta 3_k^{ts}$ is the satisfaction expectation coefficients of the customer's demands for the quality, and $\beta 4_k^{ts}$ is the satisfaction expectation coefficients of the customer's demands for the price.

Then, $\beta 1_k^{ts}, \beta 2_k^{ts}, \beta 3_k^{ts}, \beta 4_k^{ts}$ are respectively added to the corresponding satisfaction rates of Eq. 2. Therefore, Eq. 2 can be modified to Eq. 7.

$$\begin{aligned} f_1 = min \sum_{s=1}^{S}\sum_{t=1}^{T}\sum_{k=1}^{K} p_s \cdot \xi_{N1}^{ts-} \cdot \beta 1_k^{ts} + min \sum_{s=1}^{S}\sum_{t=1}^{T}\sum_{k=1}^{K} p_s \cdot \xi_{T1}^{ts-} \cdot \beta 2_k^{ts} \\ + min \sum_{s=1}^{S}\sum_{t=1}^{T}\sum_{k=1}^{K} p_s \cdot \xi_{Q1}^{ts-} \cdot \beta 3_k^{ts} + min \sum_{s=1}^{S}\sum_{t=1}^{T}\sum_{k=1}^{K} p_s \cdot \xi_{P1}^{ts-} \cdot \beta 4_k^{ts} \end{aligned} \tag{7}$$

Objective 2: Low manufacturing enterprise production cost is the objective. The model is Eq. 8 in the form:

$$f_2 = min \xi C_2^+ \tag{8}$$

$s.t.$

$$C/MC + \xi C_2^- - \xi C_2^+ = 1 \tag{9}$$

$$\begin{aligned} \sum_{s=1}^{S} p_s \sum_{t=1}^{T} \Bigg\{ \sum_{i=1}^{n} (C_{it}^s + RC_{it}^s + \sum_{j=1}^{m} C_{jt}^{'s} \cdot N_{ijt}^{'s}) \cdot \widehat{N}_{it}^s + \sum_{i=1}^{n} E_{it}^s \cdot Max(\widehat{N}_{it}^s - G_{it}^s, 0) \\ + \sum_{i=1}^{n} F_{it}^s \cdot Max(G_{it}^s + N_{it}^s, 0) \Bigg\} - C \le 0 \end{aligned} \tag{10}$$

The variables are: C is the actual cost incurred in the production process of an enterprise, MC is the ideal lowest cost incurred in the production process of an enterprise, ξC_2^- is the part that is higher than the planned cost, and ξC_2^+ is the part that is lower than the planned cost.

Objective 3: Short production cycle, short delivery time and the production arrangement are the most reasonable. The model is Eq. 11 in the form:

$$f_3 = min \sum_{s=1}^{S} \sum_{t=1}^{T} \xi T_3^{ts-} \tag{11}$$

s.t.

$$\sum_{i=1}^{N} \widehat{TN}_{it}^{s-} / TT_t^s + \xi T_3^{ts-} - \xi T_3^{ts+} = 1, \ \forall t \in N, \forall s \in S \tag{12}$$

Under the scenario *s*, TT_t^s is the minimum production cycle pursued by the *t*-th cycle production operation; ξT_3^{ts-} is a part of the delivery time demand. ξT_3^{ts+} is the exceeded part of the delivery time demand.

Objective 4: Maximization of the enterprise's profit is the objective. The model is Eq. 13 in the form:

$$f_4 = min(1 - \xi Pr_4^+) \tag{13}$$

$s.t.$

$$PRF/MPRF + \xi Pr_4^- - \xi Pr_4^+ = 1 \tag{14}$$

PRF is the actual profit produced by the enterprise; $MPRF$ is the maximum profit produced by the enterprise; ξPr_4^- and ξPr_4^+ indicate the unreached parts and the excess parts of objective 4, respectively. The equation is:

$$\begin{aligned} PRF = & \sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{i=1}^{n} (\hat{P}_{it}^s - C_{it}^s - RC_{it}^s - \sum_{j=1}^{m} C_{jt}^{'s} \cdot N_{ijt}^{'s}) \cdot \hat{N}_{it}^s \\ & - \sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{i=1}^{n} E_{it}^s \cdot Max(\hat{N}_{it}^s - G_{it}^s, 0) - \sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{i=1}^{n} F_{it}^s \cdot Max(G_{it}^s + N_{it}^s, 0) \end{aligned} \tag{15}$$

3.2 The constraints of the objective function

The objective function should satisfy the following constraints in addition to Eqs. 3, 4, 5, 6, 9, 10, 12, and 14:

$$ln_{it} = N_{it} - R_{it} \tag{16}$$

$$N_{it} - R_{it} \leq W_i \tag{17}$$

$$ln_{it} \leq W_i \tag{18}$$

$$ln_{it} \leq N_{it} \tag{19}$$

$$\sum_{i=1}^{n} N_{ijt}' - R_{jt}' \leq W_j' \tag{20}$$

$$R_{i\cdot(t+1)} = R_{i\cdot t} + N_{i\cdot(t-T_i)} - R_{i\cdot(t-T_i)} \tag{21}$$

$$R_{it} \leq MR_i, \ R_{i\cdot(t+1)} \leq MR_i \tag{22}$$

$$TN_{it} \geq T_{it} + ET_{it} \tag{23}$$

$$RT_{it} \geq ET_{it} \tag{24}$$

$$lt_{it} = T_{it} + ET_{it} - TN_{it} \tag{25}$$

$$\begin{gathered} N_{it}, \ R_{it}, \ W_{it}, \ N_{ijt}', \ R_{jt}', \ W_{jt}', \ P_{it}, \ C_{it}, \ C_{jt}', \ E_{it}, \ F_{it}, \ G_{it} \geq 0, \\ i = 1,2,\dots,n, \ j = 1,2,\dots,m, \ \forall t \in N, \forall s \in S \end{gathered} \tag{26}$$

Eqs. 16 to 20 are mainly constraints on production capacity. Eqs. 21 and 22 are mainly constraints on inventory. Eqs. 23 to 25 are mainly constraints on production time, and Eq. 26 is a nonnegatively constrained variable.

3.3 The triangular fuzzy description of the number of demands with fuzziness

For quantity demand $\widetilde{N}_{it}^{s}$, we use triangular fuzzy numbers $\widetilde{N}_{it}^{s} = (dn_{1it}^{s}, dn_{2it}^{s}, dn_{3it}^{s})$ to express the uncertain demand quantity, for which dn_{1it}^{s} and dn_{3it}^{s} are the lower and upper boundaries of the fuzzy number. That is, the number of demands is within this range $[dn_{1it}^{s}, dn_{3it}^{s}]$. dn_{2it}^{s} is the point where the degree of membership is 1. The membership function is Eq. 27 in the form:

$$\beta 1_{k}^{ts} = \rho(z) = \begin{cases} 0 & z \le dn_{1it}^{s} \\ \dfrac{(z - dn_{1it}^{s})}{(dn_{2it}^{s} - dn_{1it}^{s})}, & dn_{1it}^{s} < z < dn_{2it}^{s} \\ \dfrac{(dn_{3it}^{s} - z)}{(dn_{3it}^{s} - dn_{2it}^{s})}, & dn_{2it}^{s} < z < dn_{3it}^{s} \\ 0 & z \ge dn_{3it}^{s} \end{cases} \quad \forall t,\ s,\ i \tag{27}$$

We suppose $dn_{2it}^{s} = N_{it}^{s}$, and let $dn_{1it}^{s}, dn_{3it}^{s}$ be the upper and lower bounds of the demand quantity, we obtain the customer demand for the membership function; z is the actual demand of the customer $\widehat{N}_{it}^{s}$.

3.4 Trapezoidal fuzzy description of demand delivery with fuzziness

For the delivery time demand $\widetilde{TN}_{it}^{s}$, the research in this paper uses trapezoidal fuzzy numbers $\widetilde{TN}_{it}^{s} = (dtn_{1it}^{s}, dtn_{2it}^{s}, dtn_{3it}^{s}, dtn_{4it}^{s})$ to indicate an uncertain demand lead time. In the fuzzy numbers, dtn_{1it}^{s} and dtn_{4it}^{s} indicate the upper and lower bounds of the fuzzy number; dtn_{2it}^{s} and dtn_{4it}^{s} are the peak value of this fuzzy number. Its membership function is Eq. 28 in the form:

$$\beta 2_{k}^{ts} = \varphi(z) = \begin{cases} 0 & z \le dtn_{1it}^{s} \\ \dfrac{(z - dtn_{1it}^{s})}{(dtn_{2it}^{s} - dtn_{1it}^{s})}, & dtn_{1it}^{s} < z < dtn_{2it}^{s} \\ 1, & dtn_{2it}^{s} \le z \le dtn_{3it}^{s} \\ \dfrac{(dtn_{4it}^{s} - z)}{(dtn_{4it}^{s} - dtn_{3it}^{s})}, & dtn_{3it}^{s} < z < dtn_{4it}^{s} \\ 0 & z \ge dtn_{4it}^{s} \end{cases} \quad \forall t,\ s,\ i \tag{28}$$

3.5 Interval rough description of demand quality with roughness

For quality demand $\tilde{Q}_{it}^{s}$, we use an interval roughness of $\tilde{Q}_{it}^{s} = ([dq_{1it}^{s}, dq_{2it}^{s}], [dq_{3it}^{s}, dq_{4it}^{s}])$ representing the quality demands; the assumptions are only two roughness sets, represented by the pairs of the top approximation and the bottom approximation. Among the interval roughness variables, dq_{1it}^{s} and dq_{2it}^{s} are the lower and upper bounds of the lower approximate; dq_{3it}^{s} and dq_{4it}^{s} are the lower and upper bounds of the upper approximation. The membership function is Eq. 29 in the form:

$$\tilde{\tilde{Q}}_{it}^{s} = \omega(z) = \begin{cases} 0, & z < dq_{1it}^{s} \\ \dfrac{(dq_{1it}^{s} + dq_{2it}^{s})}{2}, & dq_{1it}^{s} \le z \le dq_{2it}^{s} \\ 0, & dq_{2it}^{s} < z < dq_{3it}^{s} \\ \dfrac{(dq_{3it}^{s} + dq_{4it}^{s})}{2}, & dq_{3it}^{s} \le z \le dq_{4it}^{s} \\ 0, & z > dq_{4it}^{s} \end{cases} \quad \forall t,\ s,\ i \tag{29}$$

3.6 Fuzzy description of fuzzy demand price

For the price demand $\tilde{P}_{it}^{s}$, because the price fluctuates within a certain range, we define a value interval, denoted as $[dp_{1it}^{s}, dp_{2it}^{s}]$, and request $dp_{1it}^{s} \le \tilde{P}_{it}^{s} \le dp_{2it}^{s}$. Among the value interval variables, dp_{1it}^{s} is the lowest price the enterprise can accept, and dp_{2it}^{s} is the highest price the cus-

tomer can accept. This adds a constraint, Eq. 30, for the price demand $\tilde{P}_{it}^{s}$. The membership function is Eq. 30 in the form:

$$dp_{1it}^{s} \le \tilde{P}_{it}^{s} \le dp_{2it}^{s} \tag{30}$$

$$\beta 4_{k}^{ts} = \begin{cases} 0 & , \quad \tilde{P}_{it}^{s} < dp_{1it}^{s} \\ \dfrac{\left(\tilde{P}_{it}^{s} - dn_{1it}^{s}\right)}{\left(dn_{2it}^{s} - dn_{1it}^{s}\right)} & , \quad dp_{1it}^{s} \le \tilde{P}_{it}^{s} \le dp_{2it}^{s} \\ 0 & , \quad \tilde{P}_{it}^{s} > dp_{2it}^{s} \end{cases} \tag{31}$$

In summary, the multi-objective production-planning model established in this paper is as follows:

- the objective function is composed of Eqs. 1, 7, 8, 11, and 13,
- the constraints are composed of Eqs. 9, 10, 12, 14, Eqs. 16 to 25, and Eqs. 27 to 31,
- Eq. 26 is the nonnegative statement.

4. Used method

The production-planning model of this paper is a multi-objective optimization problem. At the same time, it is also an NP-hard problem. It needs to be solved using a multi-objective optimization algorithm. A process analysis diagram for solving this problem, based on the NSGAII concept is presented in Fig. 2.

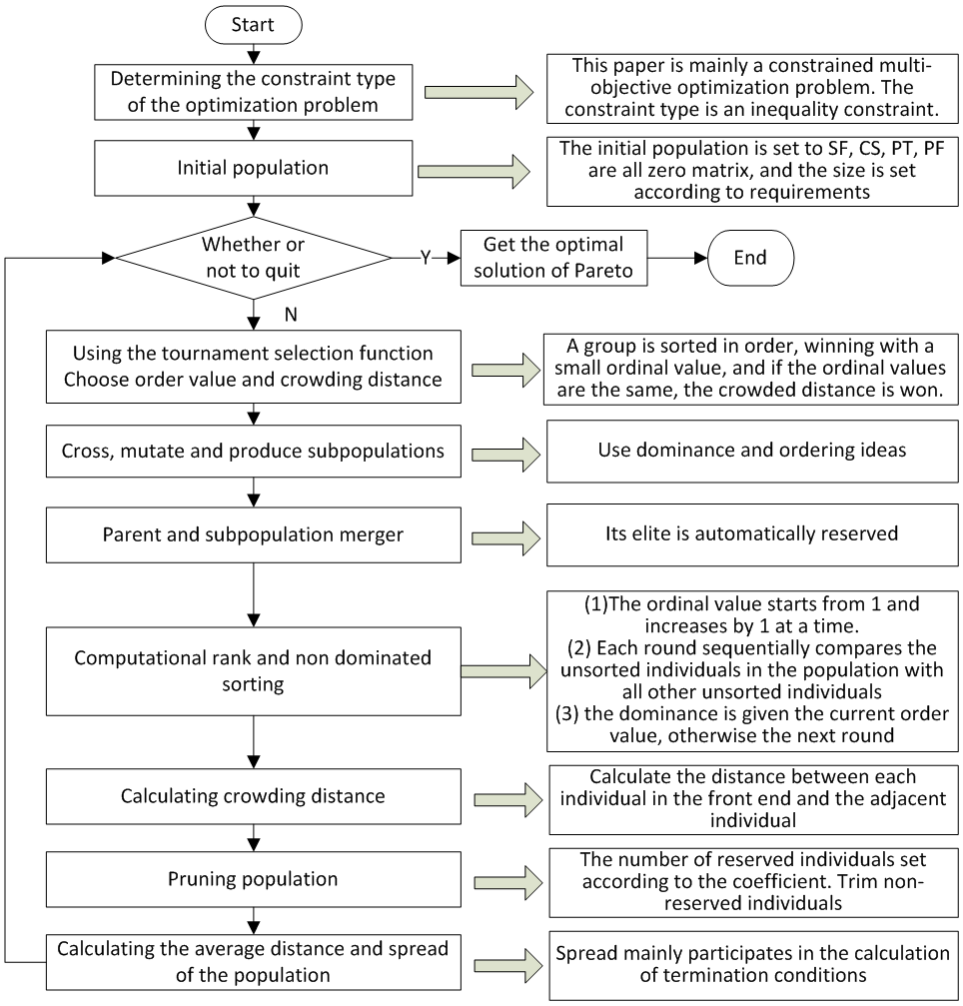


Fig. 2 Process analysis diagram for solving this problem based on the NSGAII concept

This article mainly uses Kalyanmoy Deb's Nondominated Sorting Genetic Algorithm II (NSGA-II). The NSGA-II algorithm uses an elite strategy. The next generation of populations includes a parent population and a child population. The good individuals in the parent population are preserved. Each individual group is layered and stored in two generations. Therefore, the best individuals in the group will be retained to the end. The NSGA-II algorithm is used to solve many practical problems. Fallah-Mehdipour *et al.* used this algorithm to solve a problem in building project management [18]. Huang *et al.* used this algorithm to solve a telecom customer churn problem [19]. The NSGA-II algorithm changes the fitness strategy of the sharing radius and proposes a comparison operator of congestion degree [20]. The algorithm performs fast sorting. Then, the winning criterion is set so that the individual is evenly distributed in the Pareto domain. The sorting improves calculation speed and preserves the diversity of the population.

5. Results and discussion

According to the above model, the examples are combined for simulation and analysis. MATLAB's 2015b version is adopted to carry out the simulation's calculations and analyze the results.

5.1 Initial data

(1) An equipment manufacturing enterprise is supposed to have the ability to produce 5 kinds of products, and there are 10 kinds of resources required. That is, $n = 5, m = 10$. The probability of order modification and emergency ordering is assumed to be 0.3 and 0.7, respectively. We assume that at some point $t = 1$, the initial values of some parameter variables are as follows:

$$N'_{ijt} = \begin{bmatrix} 2 & 3 & 1 & 1 & 0 & 3 & 0 & 4 & 3 & 4 \\ 0 & 4 & 4 & 2 & 0 & 3 & 3 & 0 & 4 & 4 \\ 0 & 2 & 3 & 4 & 3 & 4 & 0 & 2 & 3 & 3 \\ 1 & 3 & 0 & 0 & 2 & 3 & 1 & 0 & 0 & 4 \\ 0 & 2 & 3 & 0 & 1 & 3 & 2 & 3 & 2 & 1 \\ 0 & 3 & 3 & 0 & 4 & 1 & 2 & 1 & 0 & 3 \end{bmatrix} \qquad N_{it} = \begin{bmatrix} 1800 & 1000 & 1400 & 1000 & 800 \\ 200 & 1600 & 600 & 1600 & 1800 \\ 1600 & 2000 & 800 & 2000 & 200 \\ 2000 & 200 & 1400 & 1000 & 1800 \\ 2000 & 2000 & 400 & 1200 & 200 \end{bmatrix}$$

$$\Delta N_{it} = \begin{bmatrix} 300 & 1000 & 900 & 400 & 500 \\ 300 & 800 & 900 & 900 & 900 \\ 800 & 600 & 600 & 100 & 900 \\ 900 & 200 & 500 & 800 & 600 \\ 800 & 400 & 1000 & 900 & 300 \end{bmatrix} \qquad G_{it} = \begin{bmatrix} 5700 & 5100 & 5300 & 4700 & 5300 \\ 6500 & 5700 & 6500 & 4400 & 6000 \\ 4800 & 4100 & 6300 & 6200 & 6100 \\ 4800 & 5100 & 6600 & 7000 & 5600 \\ 4700 & 4500 & 3600 & 6400 & 5900 \end{bmatrix}$$

$$BN_{it} = \begin{bmatrix} 500 & 900 & 900 & 200 & 100 & 600 & 500 & 600 & 400 & 700 \\ 400 & 300 & 600 & 100 & 1000 & 400 & 100 & 600 & 300 & 300 \\ 500 & 900 & 100 & 1000 & 800 & 600 & 500 & 200 & 800 & 600 \\ 700 & 500 & 400 & 200 & 400 & 700 & 800 & 1000 & 700 & 200 \\ 800 & 300 & 900 & 800 & 800 & 400 & 400 & 600 & 800 & 400 \end{bmatrix}$$

$$\widehat{TN}_{it} = \begin{bmatrix} 14 & 15 & 13 & 9 & 11 & 10 & 13 & 10 & 11 & 15 \\ 8 & 10 & 10 & 6 & 11 & 10 & 11 & 12 & 11 & 12 \\ 14 & 12 & 8 & 14 & 8 & 10 & 9 & 13 & 9 & 12 \\ 10 & 15 & 15 & 10 & 12 & 15 & 11 & 11 & 12 & 10 \\ 9 & 6 & 11 & 11 & 5 & 7 & 8 & 11 & 5 & 7 \end{bmatrix}$$

$$E_{it} = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} \qquad C_{it} = \begin{bmatrix} 20 \\ 30 \\ 25 \\ 35 \\ 40 \end{bmatrix} \qquad RC_{it} = \begin{bmatrix} 10 \\ 15 \\ 12 \\ 14 \\ 12 \end{bmatrix} \qquad F_{it} = \begin{bmatrix} 5 \\ 5 \\ 5 \\ 5 \\ 5 \end{bmatrix}$$

$$T_{it} = \begin{bmatrix} 7 \\ 5 \\ 6 \\ 8 \\ 4 \end{bmatrix} \qquad RT_{it} = \begin{bmatrix} 2 \\ 1 \\ 2 \\ 2 \\ 1 \end{bmatrix} \qquad ET_{it} = \begin{bmatrix} 8 \\ 5 \\ 7 \\ 9 \\ 5 \end{bmatrix} \qquad F_{it} = \begin{bmatrix} 11 \\ 8 \\ 10 \\ 12 \\ 7 \end{bmatrix}$$

$$RC_{jt} = [2 \quad 3 \quad 2 \quad 1 \quad 2 \quad 3 \quad 4 \quad 3 \quad 1 \quad 2]'$$

(2) The triangular fuzzy number $\widetilde{BN}_{it}^{s} = (dn_{1it}^{s}, dn_{2it}^{s}, dn_{3it}^{s})$. Suppose that the triangular fuzzy number $dn_{2it}^{s} = BN_{it}$, then:

$$\widetilde{BN}_{it} =$$

$$\begin{bmatrix}
(0,500,1000) & (0,900,1000) & (0,900,1000) & (0,200,1000) & (0,100,1000) & (0,600,1000) & (0,500,1000) & (0,600,1000) & (0,400,1000) & (0,700,1000) \\
(0,400,1000) & (0,300,1000) & (0,600,1000) & (0,100,1000) & (0,1000,1000) & (0,400,1000) & (0,100,1000) & (0,600,1000) & (0,300,1000) & (0,300,1000) \\
(0,500,1000) & (0,900,1000) & (0,100,1000) & (0,1000,1000) & (0,800,1000) & (0,600,1000) & (0,500,1000) & (0,200,1000) & (0,800,1000) & (0,600,1000) \\
(0,700,1000) & (0,500,1000) & (0,400,1000) & (0,200,1000) & (0,400,1000) & (0,700,1000) & (0,800,1000) & (0,1000,1000) & (0,700,1000) & (0,200,1000) \\
(0,800,1000) & (0,300,1000) & (0,900,1000) & (0,800,1000) & (0,800,1000) & (0,400,1000) & (0,400,1000) & (0,600,1000) & (0,800,1000) & (0,400,1000)
\end{bmatrix}$$

Assuming the actual number of customer needs $\widehat{BN}_{it}$:

$$\widehat{BN}_{it} = \begin{bmatrix}
600 & 600 & 800 & 900 & 500 & 100 & 700 & 100 & 500 & 900 \\
600 & 800 & 400 & 200 & 700 & 900 & 900 & 900 & 300 & 800 \\
200 & 900 & 200 & 700 & 800 & 1000 & 100 & 400 & 200 & 300 \\
800 & 400 & 300 & 1000 & 100 & 500 & 300 & 200 & 500 & 700 \\
500 & 500 & 500 & 100 & 1000 & 500 & 500 & 800 & 200 & 100
\end{bmatrix}$$

Bring $\widehat{BN}_{it}$ as z to Eq. 27, then:

$$\beta 1_{k}^{ts} = \begin{bmatrix}
0.800 & 0.667 & 0.889 & 0.126 & 0.556 & 0.167 & 0.601 & 0.167 & 0.834 & 0.336 \\
0.667 & 0.287 & 0.667 & 0.889 & 0.700 & 0.168 & 0.112 & 0.252 & 1.000 & 0.287 \\
0.400 & 1.000 & 0.889 & 0.700 & 1.000 & 0.002 & 0.200 & 0.750 & 0.250 & 0.500 \\
0.668 & 0.800 & 0.750 & 0.001 & 0.250 & 0.714 & 0.375 & 0.200 & 0.714 & 0.376 \\
0.625 & 0.715 & 0.556 & 0.125 & 0.005 & 0.834 & 0.834 & 0.501 & 0.250 & 0.250
\end{bmatrix}$$

(3) Trapezoidal fuzzy numbers:

$$\widetilde{TN}_{it}^{s} = (dtn_{1it}^{s}, dtn_{2it}^{s}, dtn_{3it}^{s}, dtn_{4it}^{s}) = \begin{bmatrix}
(8,11,15,18) \\
(6,10,12,15) \\
(7,12,16,20) \\
(8,13,16,20) \\
(4,8,11,14)
\end{bmatrix}$$

Assume that the actual customer demand lead time $\widehat{\mathrm{TN}}_{\mathrm{it}}$ is entered as z into Eq. 28, which results in:

$$\beta 2_{k}^{ts} = \begin{bmatrix}
1.000 & 1.000 & 1.000 & 0.333 & 1.000 & 0.667 & 1.000 & 0.667 & 1.000 & 1.000 \\
0.500 & 1.000 & 1.000 & 0.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 \\
1.000 & 1.000 & 0.200 & 1.000 & 0.200 & 0.600 & 0.400 & 1.000 & 0.400 & 1.000 \\
0.400 & 1.000 & 1.000 & 0.400 & 0.800 & 1.000 & 0.600 & 0.600 & 0.800 & 0.400 \\
1.000 & 0.500 & 1.000 & 1.000 & 0.250 & 0.750 & 1.000 & 1.000 & 0.250 & 0.750
\end{bmatrix}$$

(4) Assume that each product produced by the enterprise has only two specifications (high, low), and the rough set of the quality requirements for each product is (0.3, 0.6, 0.6, 0.9). We suppose the customer quality requirements are:

$$\tilde{Q}_{it}^{s} = \begin{bmatrix}
0.900 & 0.900 & 0.600 & 0.500 & 0.600 & 0.400 & 0.400 & 0.900 & 0.700 & 0.800 \\
0.300 & 0.600 & 0.500 & 0.300 & 0.900 & 0.400 & 0.700 & 0.600 & 0.900 & 0.600 \\
0.400 & 0.300 & 0.600 & 0.900 & 0.600 & 0.700 & 0.900 & 0.800 & 0.600 & 0.500 \\
0.500 & 0.800 & 0.900 & 0.700 & 0.400 & 0.900 & 0.500 & 0.500 & 0.500 & 0.400 \\
0.600 & 0.600 & 0.600 & 0.400 & 0.400 & 0.800 & 0.900 & 0.400 & 0.300 & 0.400
\end{bmatrix}$$

and place it into z into Eq. 29, then we get:

$$\beta 3_{k}^{ts} = \begin{bmatrix}
0.750 & 0.750 & 0.450 & 0.450 & 0.450 & 0.450 & 0.450 & 0.750 & 0.750 & 0.750 \\
0.450 & 0.450 & 0.450 & 0.450 & 0.750 & 0.450 & 0.750 & 0.450 & 0.750 & 0.450 \\
0.450 & 0.450 & 0.450 & 0.750 & 0.450 & 0.750 & 0.750 & 0.750 & 0.450 & 0.450 \\
0.450 & 0.750 & 0.750 & 0.750 & 0.450 & 0.750 & 0.450 & 0.450 & 0.450 & 0.450 \\
0.450 & 0.450 & 0.450 & 0.450 & 0.450 & 0.750 & 0.750 & 0.450 & 0.450 & 0.450
\end{bmatrix}$$

(5) Assuming that the price of each product is consistent for the customer within a certain period, then as long as the price is within a certain range, it can be assumed that the customer wants the price to be $P_{it}^{s} = [100,200,150,160,140]'$,

$$\tilde{P}_{it}^{s} = (\, dp_{1it}^{s}, dp_{2it}^{s}) =$$

$$\begin{bmatrix}
(94,108) & (99,107) & (98,107) & (94,105) & (92,107) & (96,103) & (97,108) & (96,109) & (94,100) & (98,103) \\
(193,201) & (198,201) & (192,206) & (192,206) & (200,202) & (191,200) & (195,205) & (191,208) & (193,209) & (197,202) \\
(147,154) & (148,154) & (146,159) & (149,150) & (150,153) & (142,153) & (141,153) & (141,151) & (149,158) & (150,158) \\
(176,188) & (175,181) & (171,184) & (173,180) & (177,189) & (176,183) & (179,182) & (176,185) & (178,189) & (180,185) \\
(114,120) & (115,128) & (111,125) & (119,124) & (116,122) & (115,123) & (112,121) & (118,125) & (116,124) & (120,125)
\end{bmatrix}$$

Bring $\tilde{P}_{it}^{s}$ to Eq. 31

$$\beta 4_{k}^{ts} = \begin{bmatrix} 0.571 & 0.875 & 0.778 & 0.455 & 0.467 & 0.429 & 0.727 & 0.692 & 0.000 & 0.600 \\ 0.125 & 0.333 & 0.429 & 0.429 & 1.000 & 0.000 & 0.500 & 0.471 & 0.563 & 0.400 \\ 0.571 & 0.667 & 0.692 & 0.000 & 1.000 & 0.273 & 0.250 & 0.100 & 0.889 & 1.000 \\ 0.667 & 0.167 & 0.308 & 0.000 & 0.750 & 0.429 & 0.667 & 0.556 & 0.818 & 1.000 \\ 0.000 & 0.615 & 0.357 & 0.800 & 0.333 & 0.375 & 0.111 & 0.714 & 0.500 & 1.000 \end{bmatrix}$$

5.2 Operation results

We bring all the data into the formula and use MATLAB for programming a solution with the NAGAII algorithm. The objective function is a random number [1,100]. The production batch size of the enterprise setting is 100. That is, $PL = 100$. In addition, the population size of the genetic algorithm is set to 400. Its optimal front individual coefficient is set to 0.25. The algebra is 300, and the fitness function has a bias of 0.001. Fig. 3 is a diagram of the individual distances, individual average distances, and the Pareto front.

From the distribution of the running Pareto front graph, it can be seen that the solution is even. The distance of individuals indicates the distance between the individual and other individuals. It can be seen from Fig. 3 that the crowded distances differ greatly, indicating that the populations are not crowded and the population diversity is better. The average distance represents the average distance between individuals. It shows that the population distance tends to be balanced and flat. By solving the objective functions, 100 Pareto optimal solutions are obtained. Since there are four objective functions in this paper, we need to sort according to the priority of four objectives. First, the solutions are sorted in ascending order according to objective 1 (i.e., the highest customer satisfaction). In the case of the same objective 1 value, objective 2 (i.e., the lowest production cost) is sorted in ascending order. If objective 2's values are also the same, the solutions are sorted in ascending order according to objective 3 (i.e., the shortest delivery time). Finally, the solutions are sorted in descending order according to objective 4 (i.e., the largest profit). That is, objective 1>> objective 2>> objective 3>> objective 4. Fig 4 is a multi-dimensional contrast scatter plot of the first two objective functions plotted.

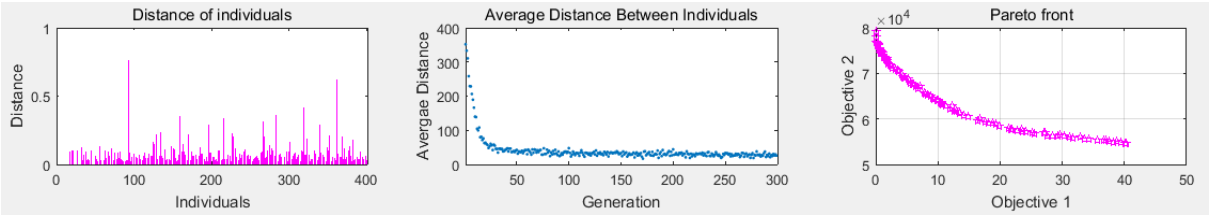


Fig. 3 The distance and the average distance between individuals and Pareto front

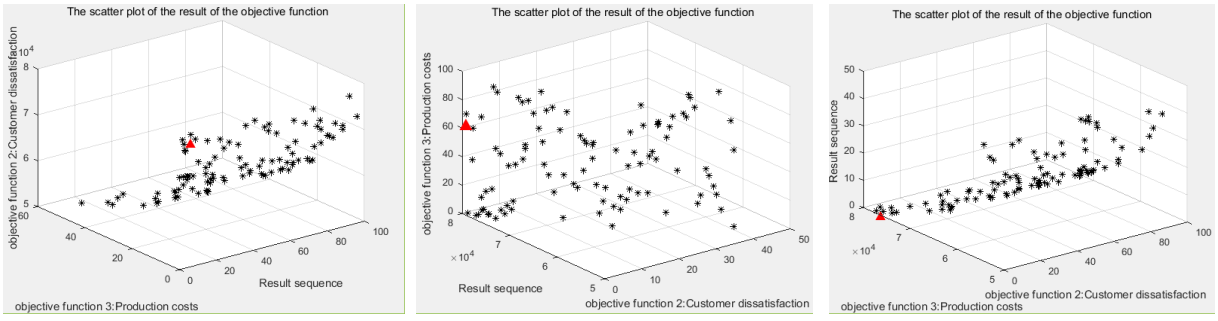


Fig. 4 Multidimensional contrast scatter plot of the first two objectives

It can be seen, from Fig. 4, that the triangle symbol in the figure indicates the point of maximum customer satisfaction, minimum production cost, and shortest delivery time. In addition, the 100 Pareto optimal solutions calculated by this model are evenly distributed in one plane. This shows that the results of the operation basically meet the results of the multi-objective

planning solution. Some of the results (first six of the 100) are shown in Table 2. From the calculation results, it can be found that the 71-st set of data satisfies the minimum customer dissatisfaction, the ratio is 0, and the minimum production cost is 79173.137. Additionally, Table 3 is the production plan quantity that runs out at time $t = 1$.

Table 2 Partial result (first six of the 75)

Customer dissatisfaction	Production cost	Production cycle	Serial number
0	79173.14	3823	71
0.06	78805.17	3734	2
0.126666667	77884.72	3409	41
0.201666667	77839.99	3486	63
0.226666667	77220.7	3328	3
0.301666667	76777.69	3185	4

Table 3 The production plan quantity ($t = 1$)

Products \ Period	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$	$k = 9$	$k = 10$	Total
$n(1)$	700	500	600	900	700	1100	800	800	600	800	7500
$n(2)$	900	800	800	900	1000	500	600	500	600	700	7300
$n(3)$	900	800	700	900	300	900	600	500	600	1000	7200
$n(4)$	1000	600	900	800	1000	500	1000	800	800	600	8000
$n(5)$	900	1000	800	800	900	800	500	800	1000	600	8100

6. Conclusion

In this paper, we analyze the various uncertainties arising from order modification and emergency orders. There were a variety of vague uncertainties in the number of customer requirements, delivery dates, quality, and prices. According to their characteristics, we use triangular fuzzy numbers, trapezoidal fuzzy numbers, fuzzy intervals, etc. to handle these multiple uncertainties. Using the scenario analysis method, two scenarios are set, namely, the emergency ordering scenario and the order modification scenario. Based on this, a multi-objective model is established to solve the production-planning problem of equipment manufacturers by using the NSGA-II algorithm. We combine the analyses of the model to simulate the model and obtain the feasibility of the model. This will help equipment manufacturing companies to rationalize their production-planning and scheduling. Future research will further refine the model and complete the dispatch of the production plan.

Acknowledgments

This work was financially supported by Liaoning Province Planning Office of Philosophy and Social Science (L18BJY029), Shenyang Planning Office of Philosophy and Social Science (18ZX019), the Liaoning Province Department of Education Project (WGD2016002), Shenyang Science and Technology Innovation Knowledge Base (SYKJ201807,201806), and Shenyang Federation Social Science Circles (Grant no. SYSK2018-05-05), the Key Program of Social Science Foundation of Liaoning Province (Grant No. L15AGL013), and the Natural Science Foundation of Liaoning Province (Grant No. 201602545). The authors wish to acknowledge the contribution of Liaoning Key Lab of Equipment Manufacturing Engineering Management, Liaoning Research Base of Equipment Manufacturing Development, Liaoning Key Research Base of Humanities and Social Sciences, Research Center of Micromanagement Theory, and Shenyang University of Technology.

References

[1] Rahdar, M., Wang, L., Hu, G. (2018). A tri-level optimization model for inventory control with uncertain demand and lead time, *International Journal of Production Economics*, Vol. 195, 96-105, doi: 10.1016/j.ijpe.2017.10.011.

[2] Ho, J.-W., Fang, C.-C. (2013). Production capacity planning for multiple products under uncertain demand conditions, *International Journal of Production Economics*, Vol. 141, No. 2, 593-604, doi: 10.1016/j.ijpe.2012.09.016.

[3] Shi, J., Zhang, G., Sha, J. (2011). Optimal production planning for a multi-product closed loop system with uncertain demand and return, *Computers & Operations Research*, Vol. 38, No. 3, 641-650, doi: 10.1016/j.cor.2010.08.008.

[4] Diabat, A., Dehghani, E., Jabbarzadeh, A. (2017). Incorporating location and inventory decisions into a supply chain design problem with uncertain demands and lead times, *Journal of Manufacturing Systems*, Vol. 43, Part 1, 139-149, doi: 10.1016/j.jmsy.2017.02.010.
[5] Pan, F., Nagi, R. (2010). Robust supply chain design under uncertain demand in agile manufacturing, *Computers & Operations Research*, Vol. 37, No. 4, 668-683, doi: 10.1016/j.cor.2009.06.017.
[6] Erdogan, S.A., Denton, B. (2013). Dynamic appointment scheduling of a stochastic server with uncertain demand, *INFORMS Journal on Computing*, Vol. 25, No. 1, 116-132, doi: 10.1287/ijoc.1110.0482.
[7] Cho, S.-H., Tang, C.-S. (2013). Advance selling in a supply chain under uncertain supply and demand, *Manufacturing & Service Operations Management*, Vol. 15, No. 2, 305-319, doi: 10.1287/msom.1120.0423.
[8] Aghezzaf, E.-H., Sitompul, C., Najid, N.M. (2010). Models for robust tactical planning in multi-stage production systems with uncertain demands, *Computers & Operations Research*, Vol. 37, No. 5, 880-889, doi: 10.1016/j.cor.2009.03.012.
[9] Chica, M., Cordón, Ó., Damas, S., Bautista, J. (2013). A robustness information and visualization model for time and space assembly line balancing under uncertain demand, *International Journal of Production Economics*, Vol. 145, No. 2, 761-772, doi: 10.1016/j.ijpe.2013.05.030.
[10] Chica, M., Bautista, J., Cordón, Ó., Damas, S. (2016). A multiobjective model and evolutionary algorithms for robust time and space assembly line balancing under uncertain demand, *Omega*, Vol. 58, 55-68, doi: 10.1016/j.omega.2015.04.003.
[11] Xu, W., Liu, L., Zhang, Q., Wang, X. (2017). A multi-object decision-making method for location model of manufacturing industry under uncertain environment, *Journal of Interdisciplinary Mathematics*, Vol. 20, No. 4, 1019-1028, doi: 10.1080/09720502.2017.1358879.
[12] Xu, W., Yu, Y., Zhang, Q. (2018). A multiobjective decision-making method of service element allocation of customer behaviour, *Discrete Dynamics in Nature and Society*, Vol. 2018, Article ID 3572094, doi: 10.1155/2018/3572094.
[13] Tang, M., Gong, D., Liu, S., Lu, X. (2017). Finding key factors affecting the locations of electric vehicle charging stations: A simulation and ANOVA approach, *International Journal of Simulation Modelling*, Vol. 16, No. 3, 541-554, doi: 10.2507/IJSIMM16(3)CO15.
[14] Tang, M., Gong, D., Liu, S., Zhang, H. (2016). Applying multi-phase particle swarm optimization to solve bulk cargo port scheduling problem, *Advances in Production Engineering & Management*, Vol. 11, No. 4, 299-310, doi: 10.14743/apem2016.4.228.
[15] Galal, N.M., El-Kilany, K.S. (2016). Sustainable agri-food supply chain with uncertain demand and lead time, *International Journal of Simulation Modelling*, Vol. 15, No. 3, 485-496, doi: 10.2507/IJSIMM15(3)8.350.
[16] Ben Said, L., Hmiden, M., Ghedira, K. (2010). A two-step transshipment model with fuzzy demands and service level constraints, *International Journal of Simulation Modelling*, Vol. 9, No. 1, 40-52, doi: 10.2507/IJSIMM09(1)4.142.
[17] Liu, Y.F., Zhang, Q.S. (2018). Solving multi-objective planning model for equipment manufacturing enterprises with dual uncertain demands using NSGA-II algorithm, *Advances in Production Engineering & Management,* Vol. 13, No. 2, 193-205, doi: 10.14743/apem2018.2.284.
[18] Fallah-Mehdipour, E., Haddad, O.B., Tabari, M.M.R., Mariño, M.A. (2012). Extraction of decision alternatives in construction management projects: Application and adaptation of NSGA-II and MOPSO, *Expert Systems with Applications*, Vol. 39, No. 3, 2794-2803, doi: 10.1016/j.eswa.2011. 08.139.
[19] Huang, B., Buckley, B., Kechadi, T.-M. (2010). Multi-objective feature selection by using NSGA-II for customer churn prediction in telecommunications, *Expert Systems with Applications*, Vol. 37, No. 5, 3638-3646, doi: 10.1016/j.eswa.2009.10.027.
[20] Tang, M., Gong, D., Liu, S., Lu, X. (2017). Finding key factors affecting the locations of electric vehicle charging stations: A simulation and ANOVA approach, *International Journal of Simulation Modelling*, Vol. 16, No. 3, 541-554, doi: 10.2507/IJSIMM16(3)CO15.

APEM journal
Advances in Production Engineering & Management
Volume 13 | Number 4 | December 2018 | pp 442–454
https://doi.org/10.14743/apem2018.4.302
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

A quantitative analysis method of greenhouse gas emission for mechanical product remanufacturing based on Petri net

Shi, J.L.[a,b,*], **Fan, S.J.**[a], **Wang, Y.J.**[a], **Cheng, J.S.**[a]

[a]Mechanical Engineering and Automation, Dalian Polytechnic University, Dalian, P.R. China
[b]Institute of Sustainable Design and Manufacturing, Dalian University of Technology, Dalian, P.R. China

ABSTRACT

The increased greenhouse gas (GHG) emission is one of the consequences of environmental change. Waste mechanical products remanufacturing is a good production mode for environment protection. Nevertheless, GHG emissions are inevitably generated in the remanufacturing system. Some uncertainties would exist in the remanufacturing system due to the different damage statuses of old mechanical products, which result in dynamic GHG emissions. Recent studies on the characteristics of GHG emissions for mechanical product remanufacturing are not yet available. This study proposed a quantitative analysis method of GHG emissions for mechanical product remanufacturing based on Petri net. In this method, the boundary of the remanufacturing is initially defined, and the dynamic characteristics of GHG emissions are analysed. Then, a GHG emission analysis model based on Petri net is constructed. Finally, the GHG emission of a PCL803 centrifugal compressor rotor remanufacturing as a case is analysed by this proposed method. This method could provide a guidance to quantitatively analyse the characteristics of GHG emissions, and suggestions for mechanical product remanufacturing to realize cleaner production and sustainability.

ARTICLE INFO

Keywords:
Mechanical product remanufacturing;
Sustainability;
Greenhouse gas emission (GHG);
Petri nets;
Resource consumption

**Corresponding author:*
shijunli0124@163.com
(Shi, J.L.)

Article history:
Received 14 November 2017
Revised 1 September 2018
Accepted 5 September 2018

1. Introduction

Climate change has become a serious environmental threat with potential impacts on global warming [1-2]. The energy and material consumptions vary in different production activities, which cause the diversified sources of greenhouse gas (GHG) emissions [3]. Therefore, it is necessary to explore an in-depth analysis of the GHG emissions for industrial activities. Comprehensive research studies have been conducted from different aspects and levels. For instance, Li *et al.* [4] constructed a carbon emission model for the manufacturing process of machine tools based on Petri net. Cao and Li [7] also proposed a simulation approach for displaying carbon emission dynamics based on hybrid Petri nets. Esteves *et al.* [6] presented a method to evaluate annually the environmental performance in terms of GHG emissions in relation to local tallow biodiesel. Teh *et al.* [7] proposed a quantified carbon footprint intensity analysis method for Australian cement and concrete production. Chen *et al.* [8] compared GHG emissions of compact fluorescent lamps with those of linear fluorescent lamps using life cycle assessment method under China's national conditions. From life cycle aspect, Murphy *et al.* [9] conducted a comprehensive, holistic evaluation of biomass-to-energy systems to reduce the production and transportation of GHG emissions by life cycle assessment (LCA) methodology. Hao *et al.* [10] compared GHG emissions and energy consumption for electric vehicle production by employing LCA

framework. Magnusson and Mácsik [11] identified significant posts for energy and GHG emissions, which were associated with the construction, use, and removal of an artificial turf field. Many other GHG emission analysis methods have also been proposed, and fruitful research results have been achieved; research has involved waste disposal [12-13], agricultural industry [14-15], and energy consumptions [16].

Remanufacturing waste mechanical products is a good production mode of resource saving, environment protection, and GHG emissions reduction. Nonetheless, GHG emissions are inevitable in the remanufacturing system. In the remanufacturing process, some uncertainties exist in the material and energy consumptions due to the different damage conditions of old products, which would cause dynamic GHG emissions. Several research have recently involved the GHG emission of product remanufacturing, for example, Tornese *et al.* [17] characterized the carbon equivalent emissions associated with pallet remanufacturing operations for two repositioning scenarios. Peng *et al.* [18] compared two types of remanufacturing cleaning technologies from the perspective of environment emissions. Bazan *et al.* [19] developed a model of GHG emissions by considering manufacturing, remanufacturing, and transportation activities with penalty tax. When considering capital and/or carbon emission constraints, Yenipazarli [20] characterized the optimal GHG emission taxation policy to deliver the benefits of product remanufacturing and maximize the total profits. To determine optimal production quantities of a new or remanufactured product, Wang *et al.* [21] examined manufacturing/remanufacturing planning issues and presented three mathematical models.

More GHG emission analysis methods about product manufacturing activity have been proposed; however, limited in-depth studies have focused on product remanufacturing system. Consequently, the greenhouse effect of the remanufacturing system cannot be deeply understood, and effective measures for resource conservation and emission reduction cannot be performed. Therefore, exploring a quantitative GHG emission analysis method for mechanical product remanufacturing is an effective way to realize greener manufacturing and sustainability.

In this study, a GHG emission analysis method for mechanical product remanufacturing based on Petri net is proposed based on the research of Li *et al.* [4] about carbon emission modelling methods for the manufacturing processes of machine tools. Moreover, the dynamic characteristics of GHG emission are comprehensively analyzed.

2. GHG emission characteristic of the mechanical product remanufacturing

2.1 GHG emission boundary of the mechanical product remanufacturing

Generally, the remanufacturing process of mechanical product includes recycling, disassembly, cleaning, inspection, repairing, and assembly, during this process, raw materials (e.g., steel, alloy, etc.) and energy (e.g., electricity, kerosene, diesel oil, etc.) would be consumed, wastes (e.g., waste solid, waste water, waste gas, etc.) are discharged, and GHG emissions are inevitably generated. The system boundary is established in view of the remanufacturing process, as shown in Fig. 1 [22]. This system boundary comprises different resource inputs (i.e., resource flow), and GHG outputs (i.e., GHG flow).

2.2 GHG emission characteristic analysis of the mechanical product remanufacturing

GHG emissions would be analyzed from the two aspects of resource production and waste discharge and disposal based on the remanufacturing process.

When old mechanical products are transported to a workshop for remanufacturing, electricity and kerosene are consumed in the disassembly stage; electricity, water, and cleaning fluid are consumed in the cleaning stage; electricity is the main energy consumed in the inspection stage; a certain amount of electricity and metal materials, such as steel, iron and alloy, are consumed in the repairing stage; and electricity is mainly consumed in the assembly and testing stages. Considerable GHG emission is generated during the resource and energy production (i.e., mining and processing), and this part of GHG emissions coming from resource production should be considered.

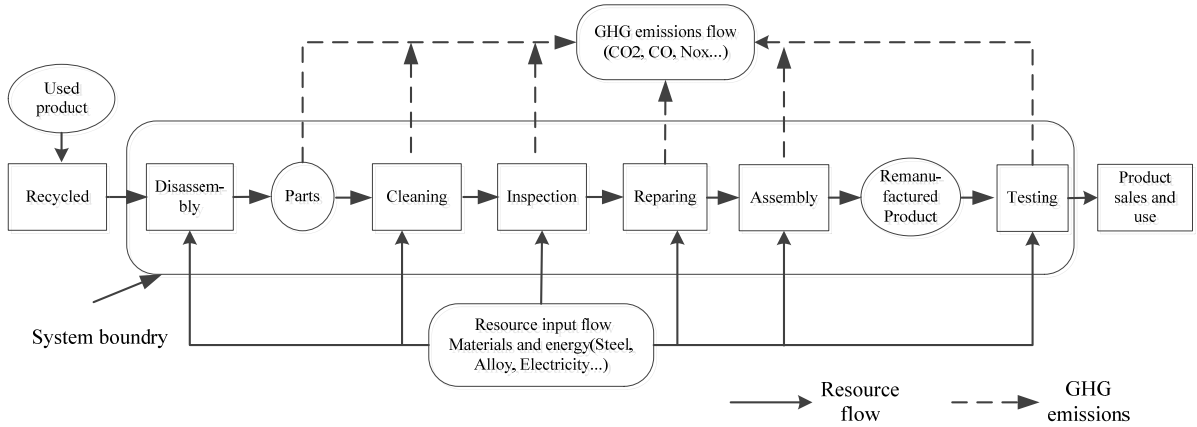


Fig. 1 GHG emissions boundary of mechanical product remanufacturing

Table 1 Main material composition of waste discharge

Waste emission	Material construction
Waste water	Cleaning water and cleaning fluid in cleaning process, the main materials are: oil, scale, carbonate, sulfate, Fe_xO_y, sodium hydroxide, sodium carbonate, alkaline solution, acid, oxidant, kerosene, etc.
Waste gas	Dust, spray powder, working gas generated in the processes of disassembly and repairing, the main materials are: Cr, Ni, Co, CO, Si, TiC, Al_2O_3, ZrO_2, SiO_2.
Solid waste	The scrap metal parts generated in the process of disassembly and the metal chip, powder and corners generated in the process of repairing.

A certain amount of waste which contains some GHG is discharged in the remanufacturing system, and in the waste disposal process, certain electricity and fossil energy are consumed, and an amount of GHG is generated during the energy production. Therefore, GHG emissions come from two sources in the waste discharge and disposal process. The main material composition of waste emissions is shown in Table 1.

3. GHG emission analysis model of the mechanical product remanufacturing

3.1 Definition of Petri net of GHG emission model

Petri net is a visual and qualitative analysis tool widely used for modelling and analysis in manufacturing system, which is composed of a continuous variable and a discrete event dynamic system. The advantages of organizational structure and dynamic behaviour allow the Petri net to clearly describe the dynamic GHG emissions by analysing the material flow, energy flow, and the dynamic situation of waste streams in the remanufacturing system. The GHG emission network model can be defined a six-element group, expressed as $\sum(P, T, F, K, W, M_0)$, where:

- $P = (P_1, P_2, P_3, \dots, P_n)$ refers to limited place sets that represent the remanufacturing process or material, semi-finished product, and finished product warehouse, which is expressed by roundness;
- $T = (T_1, T_2, T_3, \dots, T_m)$ denotes limited transition sets that represent the start or end of a remanufacturing activity, expressed by oval;
- F is a directed flow that represents the flow relation between place and transition;
- $M_0: P \to N$ is the place of initial marking;
- $K: P \to N^+ \cup \{\infty\}$ is a place capacity function, supposing the place capacity is infinite;
- $W: F \to N^+$ is a flow function, where $F = ((P \times T) \cup (T \times P))$.

A GHG emission analysis model based on Petri network can be established according to the material and energy flow, as shown in Fig. 2. The elements and constraints are listed in Table 2.

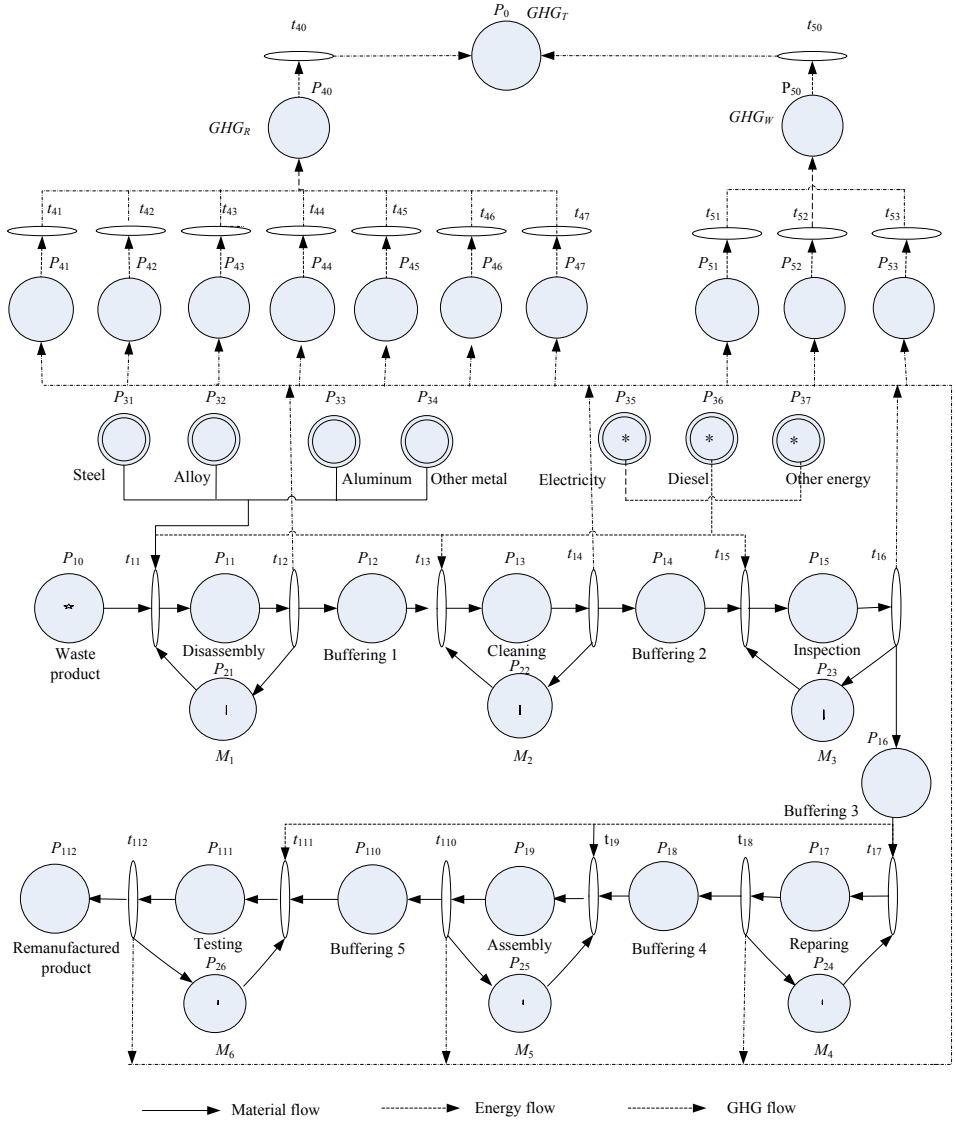


Fig. 2 GHG emissions model of mechanical product remanufacturing based on Petri network

3.2 Construction of GHG emission model

From Fig. 2, remanufacturing is associated with the consumption and transition of materials, energy, and waste. GHG emissions are generated from the following two sources: GHG emissions generated by resource production and by waste discharge and disposal. The parameter descriptions and function expressions of material, energy, and waste flows are as follows:

- i represents the i-th resource entered into the remanufacturing system, n represents the total categories of resources, the former q is material, and the latter (n-q) is energy.
- j represents the j-th remanufacturing stage, m is the total number of remanufacturing stages, k indicates the waste category, and l waste categories exist in total.

The function of the i-th resource consumption in the j-th remanufacturing stage during remanufacturing time T is expressed as follows:

$$M(i,j) = \int_0^T f_j^i(t)dt \tag{1}$$

The function of the k-th waste discharge or disposal in the j-th remanufacturing stage during remanufacturing time T is expressed as:

$$W(k,j) = \int_0^T w_j^k(t)dt \tag{2}$$

Table 2 Elements and constraints of GHG emission model

Element	Constraints explanation
P_{10}	Waste products recycled ready for remanufacturing, $M_0(P) = 1$
P_{11},P_{13},P_{15},P_{17},P_{19},P_{111}	Process of disassembly, cleaning, inspection, repairing, assembly and testing, $M_0(P) = 0$
P_{12},P_{14},P_{16},P_{18},P_{110}	Transfer process of remanufacturing , $M_0(P) = 0$
P_{21},P_{22},P_{23},P_{24},P_{25},P_{26}	Idle state of the remanufacturing process, $M_0(P) = 1$
P_{31}, P_{32}, P_{33}, P_{34}	Four major metals consumption, $M_0(P) = 0$
P_{35}, P_{36}, P_{37}	Three main energy consumption, $M_0(P) = 0$
P_{41},P_{42},P_{43},P_{44},P_{45},P_{46},P_{47}	GHG emissions generated by seven main resources (steel, electricity, …) production, $M_0(P) = 0$
P_{51},P_{52},P_{53}	GHG emissions generated by waste discharge or disposal (waste gas, water, solid), $M_0(P) = 0$
P_{40}	Total GHG emissions generated by resource production, $M_0(P) = 0$
P_{50}	Total GHG emissions generated by waste disposal, $M_0(P) = 0$
t_{ij} (i = 1,2,...,5, j = 1,2,...,12)	Start and end position of each stage
K	Capacity functions of each place of P_{ij}, $K(P_{ij}) = \infty$
W	Arc power function between the place P_{ij} and the transition t_{ij}: $W(P_{ij}, t_{ij})$ Arc weight between resource place (P_{31},P_{32},P_{33},P_{34},P_{35},P_{36},P_{37}) and transition (t_{11},t_{13},t_{15},t_{17},t_{19},t_{111}) represent the amount of resources consumption
	The arc weight function of the transition t_{ij} and the place P_{ij} is expressed as: $W(t_{ij}, P_{ij})$ The arc weights between transition t_{12},t_{14},t_{16},t_{18},t_{110},t_{112} and place P_{41}-P_{47} represent the GHG emissions generated by resources production The arc weights between transition t_{12}, t_{14}, t_{16}, t_{18}, t_{110}, t_{112} and place P_{51} ,P_{52}, P_{53} represent the GHG emissions generated by waste discharge or disposal
	All the other arc weights are 1

The parameters of GHG emissions are described as: $GHG_R^{(i,j)}$: GHG emissions generated by the i-th resource production in the j-th stage; GHG_R^i: GHG emissions generated by the i-th resource production; GHG_R: total GHG emissions generated by resource production; $GHG_W^{(k,j)}$: GHG emissions generated by the k-th waste discharge and disposal in the j-th remanufacturing stage; GHG_W^k: GHG emissions generated by the k-th waste; GHG_W: total GHG emissions generated by waste; GHG_T: total GHG emissions. The function relationship are expressed as:

$$GHG_R = \sum_{i=1}^{n} GHG_R^i = \sum_{i=1}^{n}\sum_{j=1}^{m} GHG_R^{(i,j)} \tag{3}$$

$$GHG_W = \sum_{k=1}^{l} GHG_W^k = \sum_{k=1}^{l}\sum_{j=1}^{m} GHG_W^{(k,j)} \tag{4}$$

$$GHG_T = GHG_R + GHG_W \tag{5}$$

3.3 GHG composition and characteristic parameters

GHG refers to the atmospheric gases that can absorb the solar radiation reflected by the ground that can in turn release some additional gases. The main components of GHG include CO_2, CO, NOx, and CH_4, and these gases can make the Earth's surface become warmer. Because the contribution of each gas to the potential greenhouse effect varies, the equivalent factor is used in this study to calculate the potential value of the greenhouse effect, CO_2 is considered as the benchmark gas in greenhouse effect, and its potential influence is assigned to a value of unity. The potential influences of other GHGs could be transformed to the equivalent factor of CO_2. The equivalent factors [23] of GHGs are shown in Table 3.

Table 3 Composition of greenhouse gases and equivalent factors

Environmental impact	Inventory	Equivalent Factor	Benchmark gas
GHG	CO_2	1	kg CO_2 eq
	CH_4	25	
	NOx	320	
	CO	2	

GHG emissions of resource production

Suppose that GHG_R^i represents the GHG emissions generated by the *i*-th resource production, which is expressed as

$$GHG_R^i = \sum_{c=1}^{r} M_i \times EP_c^i \tag{6}$$

where M_i represents the consumption amount of the *i*-th resource (kg), and EP_c^i is the equivalent factor (CO_2/kg ce) of the *c*-th GHG.

During remanufacturing time *T*, the GHG emissions generated by the *i*-th resource production in the *j*-th remanufacturing stage can be calculated as:

$$GHG_R^{(i,j)} = \int_0^T f_j^i(t)dt \times GHG_R^i = \sum_{c=1}^{r} \int_0^T f_j^i(t)dt \times M_i \times EP_c^i \tag{7}$$

The GHG emissions generated by the *i*-th resource production during time *T* is expressed as follows:

$$GHG_R^i = \sum_{j=1}^{m} GHG_R^{(i,j)} = \int_0^T f_j^i(t)dt \times GHG_R^i = \sum_{c=1}^{r}\sum_{j=1}^{m} \int_0^T f_j^i(t)dt \times M_i \times EP_c^i \tag{8}$$

The total GHG emissions generated by resource production during time *T* are expressed as

$$GHG_R = \sum_{i=1}^{n} GHG_R^i = \sum_{i=1}^{n}\sum_{j=1}^{m} \int_0^T f_j^i(t)dt \times GHG_R^i = \sum_{i=1}^{n}\sum_{j=1}^{k}\sum_{c=1}^{r} \int_0^T f_j^i(t)dt \times M_i \times EP_c^i \tag{9}$$

GHG emissions of waste discharge and disposal

In the waste discharge and disposal, suppose that GHG_W^k is the GHG emissions generated by the *k*-th waste discharge or disposal. Then,

$$GHG_W^k = \sum_{e=1}^{h} M_k \times EP_e^k \tag{10}$$

where M_k represents the *k*-th waste discharge or disposal, and EP_e^k is the GHG equivalent factor of the *e*-th GHG (CO_2/kg ce).

The GHG emissions generated by the *k*-th waste discharge or disposal in the *j*-th remanufacturing stage during time *T* are calculated as follows:

$$GHG_W^{(k,j)} = \int_0^T w_j^k(t)dt \times GHG_W^k = \sum_{e=1}^{h} \int_0^T w_j^k(t)dt \times M_k \times EP_e^k \tag{11}$$

The GHG emissions generated by the *k*-th waste discharge or disposal during time *T* are expressed as:

$$GHG_W^k = \sum_{j=1}^{m} GHG_W^{(k,j)} = \int_0^T w_j^i(t)dt \times GHG_W^k = \sum_{j=1}^{m}\sum_{e=1}^{h} \int_0^T w_j^k(t)dt \times M_k \times EP_e^k \tag{12}$$

The total GHG emissions generated by waste discharge and disposal during time *T* can be expressed as:

$$GHG_W = \sum_{i=1}^{l} GHG_W^k = \sum_{k=1}^{l}\sum_{j=1}^{m} \int_0^T w_j^k(t)dt \times GHG_W^k = \sum_{k=1}^{l}\sum_{j=1}^{m}\sum_{e=1}^{h} \int_0^T w_j^k(t)dt \times M_k \times EP_e^k \tag{13}$$

4. Case study

PCL803 centrifugal compressor is commonly used in long distance pipeline unit, which serves the key project of "west-east gas transmission" in China. The compressor rotor, as the core part of the equipment is with a complex manufacturing process and highly manufacturing cost. Therefore, it is of high economic value to remanufacture the compressor rotor. The object in this study is a recycled old PCL803Centrifugal compressor rotor remanufactured by a well-known large compressor manufacturer in China. The GHG emissions are analyzed in the following section.

4.1 Resource consumption and waste discharge and disposal analyses

Alloy steel, stainless steel, carbon steel and babbitt alloy are the main metal materials consumed in compressor rotor remanufacturing. The *i*-th material consumption function is expressed as follows:

$$M_i = K_i \times T \tag{14}$$

where K_i represents the average consumption of the *i*-th metal each day (kg/d), and T is the remanufacturing time (d).

In compressor rotor remanufacturing, the consumed energy mainly includes electricity, diesel, and kerosene. The relationship between energy consumption M_e and remanufacturing time T presents a linear correlation, the electricity consumption function is expressed as

$$M_e = 24 \times P \times T \tag{15}$$

where P is the average power of electricity (kW).

The consumption function of kerosene and diesel is calculated as follows:

$$M_f = V_f \times T \tag{16}$$

where V_f is the consumption rate of kerosene or diesel (kg/d).

According to an on-the-spot investigation and Eq. 14-16, the resource consumption functions in each compressor rotor remanufacturing stage are established, as shown in Table 4.

Wastes are mainly waste water, waste gases, and waste metals, and the function relation of waste water and remanufacturing time T is

$$W_w = V_w \times T \tag{17}$$

where V_w is the quantity of waste water discharge (kg/d).

The function relation of waste gases and remanufacturing time T is

$$W_g = V_g \times T \tag{18}$$

where V_g is the quantity of waste gas discharge (m^3/d).

The function relation of waste metals and remanufacturing time T is

$$W_s = V_s \times T \tag{19}$$

where V_s is the quantity of waste metal (kg/d).

Table 4 Resource consumption function in each compressor rotor remanufacturing process $M(i,j)$

Resource consumption	Compressor rotor remanufacturing stage					
	Disassembly	Cleaning	Inspection	Repairing	Assembly	Testing
Alloy steel $M_a = K_a \times T$	-	-	-	336 T	-	-
Stainless steel $M_s = K_s \times T$	-	-	-	63.15 T	-	-
Carbon steel $M_c = K_c \times T$	-	-	-	76.95 T	-	-
Bobbitt alloy $M_b = K_b \times T$	-	-	-	36 T	-	-
Kerosene $M_k = V_k \times T$	7.2 T	36 T	-	-	-	-
Diesel $M_d = V_d \times T$	30 T	99 T	-	-	-	-
Electricity $M_e = 24 \times P \times T$	45 T	222 T	93 T	30 T	18 T	15 T

Table 5 Waste discharge functions in each compressor rotor remanufacturing process $W(k,j)$

Waste emission	Compressor rotor remanufacturing stage					
	Disassembly	Cleaning	Inspection	Repairing	Assembly	Testing
Waste water $W_w = V_w \times T$	-	7.23 T	-	-	-	-
Waste gas $W_g = V_g \times T$	210 T	*19.5 T*	-	16.2 T	75 T	-
Waste solid $W_s = V_s \times T$	1536 T	10.2 T	1.5 T	2.4 T	-	-

The waste discharge functions of compressor rotor remanufacturing can be established according to Eq. 17-19, as shown in Table 5.

4.2 GHG emission calculation based on Petri net

The compressor rotor remanufacturing time in each stage comprises: 78 h of disassembly, 85 h of cleaning, 50 h of inspection, 240 h of repairing, 46 h of assembly, and 80 h of testing. The simulation benchmark is 1 d in the Petri net model, and the resource consumption and GHG emission simulation cycle is 365 d. The parameters of the Petri net model are set as follows:

Time delay parameter setting

The time delay of the remanufacturing process place is expressed as

$$D_i = (process\ time \times 2)/24 \tag{20}$$

From Eq. 20, the time delays of each remanufacturing place for P_{11}, P_{13}, P_{15}, P_{17}, P_{19}, and P_{111} are D_{11} = 6.5 d, D_{13} = 7.08 d, D_{15} = 4.17 d, D_{17} = 20 d, D_{19} = 3.83 d, D_{111} = 6.67 d.

The other time delay of the remanufacturing process place is 0.

Initial place mark setting

The place mark of compressor rotor remanufacturing is the resource consumption of 1 year (365 d). Resource consumption function $G(i, j)$ in Table 4 indicates that the initial marks of resources supply places of P_{31}-P_{37} are:

$M_0(P_{31})$ = 122640 kg, $M_0(P_{32})$ = 56994.75 kg, $M_0(P_{33})$ = 28086.75 kg
$M_0(P_{34})$ = 13140 kg, $M_0(P_{35})$ = 15768 kg, $M_0(P_{36})$ = 47085 kg, $M_0(P_{37})$ = 154396 kW

The initial mark of the other place is 0.

Place capacity

The capacity of each place P is assumed to be infinite.

Arc weight function setting

The arch weights between places P_{31}–P_{37} and transitions t_{11}, t_{13}, t_{15}, t_{17}, t_{19}, and t_{110} represent the resource consumption in each stage. The production of remanufactured compressor rotor is 1000 per year and average 3 per day. According to Table 4, the arc weight W (P_i, t_j) is determined, the results are as shown in Table 6. The arc weights between transitions t_{12}, t_{14}, t_{16}, t_{18}, t_{110}, t_{112} and places P_{41}-P_{47} represent the GHG emissions generated by the production of alloy steel, stainless steel, carbon steel, babbitt alloy, kerosene, diesel, and electricity respectively. The GHG emissions generated by resource production are referred to Chinese Life Cycle Database (CLCD) [24], shown in Table 7. The GHG emission weight by resource production could be determined according to Eq. 6-9, as shown in Table 8.

The arc weight between transitions t_{12}, t_{14}, t_{16}, t_{18}, t_{110}, t_{112} and places P_{51}-P_{53} represent the GHG emissions generated by waste discharge or disposal. The average electricity consumption of waste metal disposal is 0.45 kWh/kg. For waste water disposal, the consumed energy is mainly electricity, and the average consumption is 0.3 kWh/m^3. In the remanufacturing process, the GHG emission is mainly CO, and the density is 1.25 kg/m^3, assuming that 1 % of waste gases discharge is CO. According to Eq. 6, GHG emissions generated by waste discharge and disposal are obtained, shown in Table 9. Based on Eq. 10-13, Tables 5 and 9, GHG emission arc weights generated by waste discharge and disposal are determined, as shown in Table 10.

Table 6 Weight of resource consumption $W(P_i, t_j)$

	t_{11}	t_{13}	t_{15}	t_{17}	t_{19}	t_{111}
P_{31}	-	-	-	$W(P_{31},t_{17}) = 112$	-	-
P_{32}	-	-	-	$W(P_{32},t_{17}) = 21.5$	-	-
P_{33}	-	-	-	$W(P_{33},t_{17}) = 25.65$	-	-
P_{34}	-	-	-	$W(P_{34},t_{17}) = 12$	-	-
P_{35}	$W(P_{35},t_{11}) = 2.4$	$W(P_{35},t_{13}) = 1.2$	-	-	-	-
P_{36}	$W(P_{36},t_{11}) = 10$	$W(P_{36},t_{13}) = 33$	-	-	-	-
P_{37}	$W(P_{37},t_{11}) = 15$	$W(P_{37},t_{13}) = 74$	$W(P_{37},t_{15}) = 31$	$W(P_{34},t_{17}) = 10$	$W(P_{34},t_{17}) = 6$	$W(P_{34},t_{17}) = 5$

Table 7 GHG emission per unit resource production (kg)

GHG	Equivalent Factor	Alloy steel	Stainless steel	Carbon steel	Babbitt alloy	Kerosene	Diesel	Electricity
CO_2	1	2.75E+00	4.99E+00	2.03E+00	1.9E+00	1.9E-02	4.0E-04	9.1E-01
CO	2	2.05E-02	2.04E-02	2.99E-02	3.3E-02	2.5E+01	3.8E-01	2.0E-04
NOx	320	5.28E-03	1.24E-02	2.52E-03	5.8E-03	1.0E-01	6.0E-04	2.6E-03
CH_4	25	4.67E-03	1.28E-02	4.25E-03	5.8E-03	7.2E-02	2.1E-02	2.7E-03
GHGs of the *i*-th resources		4.60E+00	9.32E+00	3.00E+00	4.0E+00	1.5E+00	8.4E+01	1.8E+00

Table 8 GHG emission weight generated by resource production $W(t_j, P_i)$

	P_{41}	P_{42}	P_{43}	P_{44}	P_{45}	P_{46}	P_{47}
t_{12}	-	-	-	-	$W(t_{12},P_{45}) = 3.6$	$W(t_{12},P_{46}) = 18$	$W(t_{12},P_{47}) = 81$
t_{14}	-	-	-	-	$W(t_{14},P_{45}) = 1.8$	$W(t_{14},P_{46}) = 59.4$	$W(t_{14},P_{47}) = 399.6$
t_{16}	-	-	-	-	-	-	$W(t_{16},P_{47}) = 167.4$
t_{18}	$W(t_{18},P_{41}) = 514$	$W(t_{18},P_{42}) = 196.1$	$W(t_{18},P_{43}) = 77$	$W(t_{18},P_{44}) = 48$	-	-	$W(t_{18},P_{47}) = 54$
t_{110}	-	-	-	-	-	-	$W(t_{110},P_{47}) = 32.4$
t_{112}	-	-	-	-	-	-	$W(t_{112},P_{47}) = 27$

Table 9 Waste quantity and GHG emission generated by waste disposal

Waste	Energy consumption of waste disposal			GHG kg CO_2 ce
	Waste quantity	Electricity consumption	CO_2 Equivalent factor	
Waste water	2.41 m^3	Electricity 0.72 kWh	1.8kg CO_2 ce/kWh	1.3
Waste gas	106.9m^3	CO 1.3 kg	2 kg CO_2 ce	2.6
Waste solid	516.7 kg	Electricity 232.5 kWh	1.8kg CO_2 ce/kWh	418.5

Table 10 GHG emission weight generated by waste discharge and disposal $W(t_j, P_k)$

	P_{51}	P_{52}	P_{53}
t_{12}	-	$W(t_{12}, P_{52}) = 1.70$	$W(t_{12}, P_{53}) = 412$
t_{14}	$W(t_{14}, P_{51}) = 1.3$	$W(t_{14}, P_{52}) = 0.16$	$W(t_{14}, P_{53}) = 2.75$
t_{16}	-	-	$W(t_{16}, P_{53}) = 5.56$
t_{18}	-	$W(t_{18}, P_{52}) = 0.13$	$W(t_{18}, P_{53}) = 0.65$
t_{110}	-	$W(t_{110}, P_{52}) = 0.61$	-
t_{112}	-	-	-

4.3 GHG emission results

By Table 8 and Table 10, GHG emissions generated by above two sources are calculated according to Eqs. 1-5, as follows:

GHG emission place generated by resource production (unit: kg CO_2):

$$M(P_{41}) = 514,\ M(P_{42}) = 196.1,\ M(P_{43}) = 77,\ M(P_{44}) = 48$$
$$M(P_{45}) = 5.4,\ M(P_{46}) = 77.4,\ M(P_{47}) = 761.4,\ M(P_{40}) = 1679.3$$

GHG emission place generated by waste discharge and disposal (unit: kg CO_2):

$$M(P_{51}) = 1.3,\ M(P_{52}) = 2.6,\ M(P_{53}) = 420.96,\ M(P_{50}) = 424.9$$

Total GHG emission place (unit: kg CO_2):

$$M(P_0) = M(P_{40}) + M(P_{50}) = 2104.2$$

4.4 Result analysis

The GHG emissions in different remanufacturing stage are shown in Table 11 and Fig. 3, and the comparison of GHG emissions generated by two sources is shown in Fig. 4 (expressed in logarithmic form).

Table 11 Total GHG emissions in each compressor rotor remanufacturing stage (kg)

GHG emissions category	GHG emissions (CO_2 eq kg)						Total
	Disassembly	Cleaning	Inspection	Repairing	Assembly	Testing	
Alloy steel	-	-	-	514	-	-	514
Stainless steel	-	-	-	196.1	-	-	196.1
Carbon steel	-	-	-	77	-	-	77
Babbitt alloy	-	-	-	48	-	-	48
Kerosene	3.6	1.8	-	-	-	-	5.4
Diesel	18	59.4	-	-	-	-	77.4
Electricity	81	399.6	167.4	54	32.4	27	761.4
GHG_R total	102.6	460.8	167.4	889.1	32.4	27	1679.3
Waste water	-	1.3	-	-	-	-	1.3
Waste gas	1.7	0.16	-	0.13	0.61	-	2.6
Waste solid	412	2.75	5.56	0.65	-	-	420.96
GHG_W total	413.7	4.21	5.6	0.78	0.61	0	424.9
GHG_T total	516.3	465	173	889.88	33.41	27	2104.2

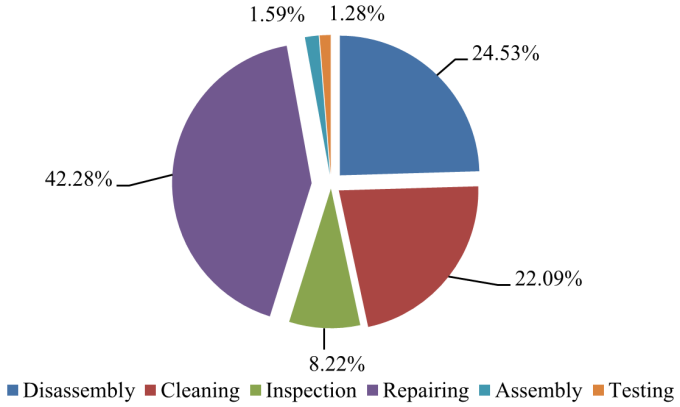


Fig. 3. GHG emission proportion in each rotor remanufacturing stage

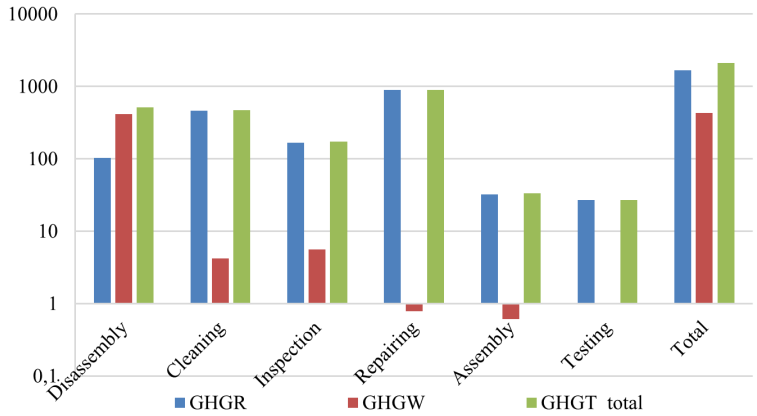


Fig. 4 GHG emission comparison of resource consumption and wastes discharge for rotor remanufacturing

GHG emission analysis in each remanufacturing stage

Table 11, Fig. 3 and Fig. 4 reveal that the most GHG emissions are generated in the repairing stage (42.28 % of the total GHG emissions), followed by those in the disassembly stage (24.53 % of the total), and the least GHG emissions are generated in the testing stage (1.28 % of the total).

Table 11 and Fig. 4 show that in resource production, most GHG emissions are produced, about 79.80 % of the entire GHG emissions (1679.3/2104.2 × 100 % = 79.80 %), especially in repairing, cleaning and inspection stages, accounting for 90.25 % of the total GHG emissions generated by resource production ((889.1 + 460.8 + 167.4)/1679.3 × 100 % = 90.25 %). For waste discharge and disposal, the GHG emissions in the disassembly stage are more than those in other stages, accounting for 97.36 % of the total GHG emissions generated by waste discharge and disposal (413.7/424.9 × 100 % = 97.36 %).

Above results are caused by the significant consumption of alloy steel, stainless steel, as well as electricity in repairing and cleaning stages. On the other hand, because most waste metals are discharged in the assembly stage and significant amounts of electricity are consumed to dispose these wastes. Therefore, large amounts of GHG emission are generated during these resource production. By contrast, almost no waste discharge and minimal electricity consumption in the testing and assembly stages, which resulted fewer GHG emissions.

GHG emission analysis of resource production

Table 11 implies that most GHG emissions are generated from electricity production in resource production, which account for 50.74 % (761.4/1679.3 × 100 % = 50.74 %) of the total GHG emissions generated by resource production; followed by alloy steel and stainless steel production, which are 30.61 % (514/1679.3 × 100 % = 30.61 %) and 11.67 % (196.1/1679.3 × 100 % = 11.67 %) respectively; and the GHG emissions of kerosene production are the least, which account only for 0.32 % (5.4/1679.3 × 100 % = 0.32 %) of the total GHG emissions.

GHG emission analysis of waste discharge and disposal

Table 11 presents that solid waste disposal produce large GHG emissions, which account for 99.05 % (420.96/424.9 × 100% = 99.05 %) of the total GHG emission generated by waste discharge and disposal, because considerable electricity is consumed. The proportion of GHG emissions generated by waste water and gas discharge (0.95 %) is only small, but they cannot be ignored.

4.5 Discussions on countermeasures

Compressor rotor remanufacturing greatly reduced GHG emissions, however, further efforts are required to improve the remanufacturing process. First, reducing the consumption of metal materials (e.g., alloy steel and stainless steel) that generate high GHG emissions in the production process is the most important approach, alternative materials that generate minimal GHG emissions are encouraged to be utilized in the compressor rotor repairing stage, therefore, optimization of process route, methods, equipment, parameters, and scheme is necessary in remanufacturing process design. Second, reduction energy consumption (mainly electricity) in the compressor rotor cleaning and inspection stage is also an effective way to decrease GHG emissions. Finally, waste metal production should be minimized in the disassembly stage, disassembly efficiency should be improved, and old compressor rotor parts should be fully recycled and remanufactured.

5. Conclusion

This study presented a quantitative GHG emission analysis model based on Petri net for mechanical product remanufacturing. In this model, the remanufacturing system boundary of mechanical products is initially determined, and GHG emission characteristics are then analyzed. Subsequently, the GHG emission model based on Petri net is constructed, and dynamic GHG emissions are analyzed according to the dynamic resource consumption and waste discharge in different remanufacturing stages. Finally, GHG emission analysis for an old PCL803 centrifugal compressor rotor remanufacturing is conducted. This model provides an effective way to analyze the GHG emissions for mechanical product remanufacturing, and the countermeasures for GHG emission reduction are accordingly introduced. The results of this typical application provide a

reference and guidance for cleaner production in product remanufacturing industry and resource selection for greener consumption and sustainability.

In fact, this method could also be applied to other mechanical product remanufacturing and newly manufacturing. For product newly manufacturing when employ this GHG analysis method, the boundary would be modified in accordance with the manufacturing process. Simplifications and hypotheses may also increase the uncertainties for other future cases in results, therefore the sensitivity analysis should be performed when utilizing this proposed model. Computer-aided software could be further developed to achieve the generality and wide applications of this method.

In addition, this study only conducted the GHG emission analysis for mechanical product remanufacturing, other types of environmental impacts, such as acidification, eutrophication, and human toxicity, are not involved. Harmful substances of SO_2, CFCs, COD, and smoke dust are inevitably produced whether in the resource production or remanufacturing process, the analysis methods of the environmental impacts generated by these harmful emissions need to be further researched. Furthermore, average data of GHG emissions are used in this study, and supplementary research for more detailed dynamics of resource consumption should be conducted in the future.

Acknowledgement

The authors gratefully acknowledge the support of Jinan Fuqiang power Co., LTD and Liaoning Province Natural Science Foundation (Energy efficiency, emissions and environmental impact model research of equipment remanufacturing system, No. 20170540080). The authors would like to thank the editor and reviewers for their constructive suggestions of the paper.

References

[1] Yang, L., Wang, J., Shi, J. (2017). Can China meet its 2020 economic growth and carbon emissions reduction targets?, *Journal of Cleaner production*, Vol. 142, Part 2, 993-1001, doi: 10.1016/j.jclepro.2016.08.018.

[2] Roelfsema, M., den Elzen, M., Höhne, N., Hof, A.F., Braun, N., Fekete, H., Böttcher, H., Brandsma, R., Larkin, J. (2014). Are major economies on track to achieve their pledges for 2020? An assessment of domestic climate and energy policies, *Energy Policy*, Vol. 67, 781-796, doi: 10.1016/j.enpol.2013.11.055.

[3] Pegallapati, A.K., Frank, E.D. (2016). Energy use and greenhouse gas emissions from an algae fractionation process for producing renewable diesel, *Algal Research*, Vol. 18, 235-240, doi: 10.1016/j.algal.2016.06.019.

[4] Li, X.-G., Li, C.-B., Liu, F., Li, L.-L. (2012). Modelling and quantification methods for carbon emission in machine tools manufacturing processes based on Petri nets, *Computer Integrated Manufacturing System*, Vol. 18, No. 2, 723-735, doi: 10.13196/j.cims.2012.12.147.lixg.005.

[5] Cao, H., Li, H. (2014). Simulation-based approach to modeling the carbon emissions dynamic characteristics of manufacturing system considering disturbances, *Journal of Cleaner Production*, Vol. 64, 572-580, doi: 10.1016/j.jclepro.2013.10.002.

[6] Esteves, V.P.P., Esteves, E.M.M., Bungenstab, D.J., Feijó, G.L.D., Araújo, O.d.Q.F., Morgado, C.d.R.V. (2017). Assessment of greenhouse gases (GHG) emissions from the tallow biodiesel production chain including land use change (LUC), *Journal of Cleaner Production*, Vol. 151, 578-591, doi: 10.1016/j.jclepro.2017.03.063.

[7] Teh, S.H., Wiedmann, T., Castel, A., de Burgh, J. (2017). Hybrid life cycle assessment of greenhouse gas emissions from cement, concrete and geopolymer concrete in Australia, *Journal of Cleaner Production*, Vol. 152, 312-320, doi: 10.1016/j.jclepro.2017.03.122.

[8] Chen, S., Zhang, J., Kim, J. (2017). Life cycle analysis of greenhouse gas emissions for fluorescent lamps in mainland China, *Science of the Total Environment,* Vol. 575, 467-473, doi: 10.1016/j.scitotenv.2016.07.058.

[9] Murphy, F., Sosa, A., McDonnell, K., Devlin, G. (2016). Life cycle assessment of biomass-to-energy systems in Ireland modelled with biomass supply chain optimisation based on greenhouse gas emission reduction, *Energy*, Vol. 109, 1040-1055, doi: 10.1016/j.energy.2016.04.125.

[10] Hao, H., Qiao, Q., Liu, Z., Zhao, F. (2017). Impact of recycling on energy consumption and greenhouse gas emissions from electric vehicle production: The China 2025 case, *Resources, Conservation and Recycling,* Vol. 122, 114-125, doi: 10.1016/j.resconrec.2017.02.005.

[11] Magnusson, S., Mácsik, J. (2017). Analysis of energy use and emissions of greenhouse gases, metals and organic substances from construction materials used for artificial turf, *Resources, Conservation and Recycling*, Vol. 122, 362-372, doi: 10.1016/j.resconrec.2017.03.007.

[12] Zeng, S., Chen, X., Dong, X., Liu, Y. (2017). Efficiency assessment of urban wastewater treatment plants in China: Considering greenhouse gas emissions, *Resources, Conservation and Recycling*, Vol. 120, 157-165, doi: 10.1016/j.resconrec.2016.12.005.

[13] Liu, Y., Ni, Z., Kong, X., Liu, J. (2017). Greenhouse gas emissions from municipal solid waste with a high organic fraction under different management scenarios, *Journal of Cleaner Production*, Vol. 147, 451-457, doi: 10.1016/j.jclepro.2017.01.135.
[14] Garcia, J.C.C., Von Sperling, E. (2017). Greenhouse gas emissions from sugar cane ethanol: Estimate considering current different production scenarios in Minas Gerais, Brazil, *Renewable and Sustainable Energy Reviews*, Vol. 72, 1033-1049, doi: 10.1016/j.rser.2017.01.046.
[15] Nunes, F.A., Seferin, M., Maciel, V.G. Flôres, S.H., Ayub, M.A.Z. (2016). Life cycle greenhouse gas emissions from rice production systems in Brazil: A comparison between minimal tillage and organic farming, *Journal of Cleaner Production*, Vol. 139, 799-809, doi: 10.1016/j.jclepro.2016.08.106.
[16] Squalli, J. (2017). Renewable energy, coal as a baseload power source, and greenhouse gas emissions: Evidence from U.S. state-level data, *Energy,* Vol. 127, 479-488, doi: 10.1016/j.energy.2017.03.156.
[17] Tornese, F., Carrano, A.L, Thorn, B.K., Pazour, J.A., Roy, D. (2016). Carbon footprint analysis of pallet remanufacturing, *Journal of Cleaner Production*, Vol. 126, 630-642, doi: 10.1016/j.jclepro.2016.03.009.
[18] Peng, S., Li, T., Tang, Z., Shi, J., Zhang, H. (2016). Comparative life cycle assessment of remanufacturing cleaning technologies, *Journal of Cleaner Production,* Vol. 137, 475-489, doi: 10.1016/j.jclepro.2016.07.120.
[19] Bazan, E., Jaber, M.Y., El Saadany, A.M.A. (2015). Carbon emissions and energy effects on manufacturing-remanufacturing inventory models, *Computers & Industrial Engineering,* Vol. 88, 307-316, doi: 10.1016/j.cie.2015.07.002.
[20] Yenipazarli, A. (2016). Managing new and remanufactured products to mitigate environmental damage under emissions regulation, *European Journal of Operational Research*, Vol. 249, No. 1, 117-130, doi: 10.1016/j.ejor.2015.08.020.
[21] Wang, Y., Chen, W., Liu, B. (2017). Manufacturing/remanufacturing decisions for a capital-constrained manufacturer considering carbon emission cap and trade, *Journal of Cleaner Production*, Vol. 140, Part 3, 1118-1128, doi: 10.1016/j.jclepro.2016.10.058.
[22] Shi, J., Li, T., Peng, S., Liu, Z., Zhang, H., Jiang, Q. (2015). Comparative life cycle assessment of remanufactured liquefied natural gas and diesel engines in China*, Journal of Cleaner Production*, Vol. 101, 129-136, doi: 10.1016/j.jclepro.2015.03.080.
[23] Guinee, J.B. (2002). *Handbook on life cycle assessment: Operational guide to the ISO standards,* Kluwer Academic Publishers, Dordrecht, Netherlands, doi: 10.1007/BF02978897.
[24] Liu, X., Wang, H., Chen, J., He, Q., Zhang, H., Jiang, R., Chen, X., Hou, P. (2010). Method and basic model for development of Chinese reference life cycle database, *Acta Scientiae Circumstantiae*, Vol. 30, No. 10, 2136-2144, doi: 10.13671/j.hjkxxb.2010.10.028.

APEM journal
Advances in Production Engineering & Management
Volume 13 | Number 4 | December 2018 | pp 455–465
https://doi.org/10.14743/apem2018.4.303
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Comprehensive analysis and study of the machinability of a high strength aluminum alloy (EN AW-AlZn5.5MgCu) in the high-feed milling

Duplák, J.[a], **Hatala, M.**[a], **Dupláková, D.**[a,*], **Steranka, J.**[b]

[a]Technical University of Košice, Faculty of Manufacturing Technologies with a seat in Prešov, Department of Automobile and Manufacturing Technologies, Prešov, Slovakia
[b]Technical University of Košice, Faculty of Manufacturing Technologies with a seat in Prešov, Department of Computer Aided Manufacturing Technologies, Prešov, Slovakia

ABSTRACT

This article is focused on studying and analysing the efficiency of the machinability of a high strength aluminium alloy (EN AW-AlZn5.5MgCu) in the high feed milling. The introduction of the article provides a brief description of high feed milling technology and presents best known research regards to the subject. The research of the time efficiency and economic efficiency of high feed milling of aluminium alloys consists of realization of two groups of experiments. The first group consists of four experiments carried out by progressive technology of high feed milling, and the second group contains one experiment conducted using conventional milling technology. The assessment of efficiency consists in determining the overall time and economic efficiency and also in comparison to the machining of aluminium alloys by high feed milling technology with conventional machining technology. The best results were obtained when the machining parameters were: cutting speed of 550 m/min, travel speed of 10600 mm/min, and feed per tooth of 0.85 mm. The material was removed in the contour roughing phase with a 42 mm plunge-cutting router. Using this cutter, it is possible to produce 19 pieces of components in a hour, which is more than half of the specified requirement. The production of components under the conditions and with this type of high feed milling cutter is more than 75 % shorter than production by a conventional method.

ARTICLE INFO

Keywords:
High-feed milling;
High strength aluminum alloy (EN AW-AlZn5.5MgCu);
Machinability; Efficiency;
Optimization

**Corresponding author:*
darina.duplakova@tuke.sk
(Dupláková, D.)

Article history:
Received 9 July 2018
Revised 24 August 2018
Accepted 27 August 2018

1. Introduction

Comprehensive simulation, exploration and analysis of the efficiency of the production process using high feed milling as a part of the major targets of the workload of current scientific practice. High feed machining determines the overall machining efficiency as the feed rate is equal to multiplying the feed by speed. Examining machinability, optimizing cutting parameters and all the influencing parameters, is immensely important to the efficient economic and cost-effective adjustment of the production process [1-3]. Competition in the global market in cutting operations is increasing from one hour to the next. Technologies, methods and processes to produce more efficiently and cheaper are daily innovated. A fundamental principle of each company is to achieve the minimum production time in connection with the minimization of production costs,

thus ensuring an efficient production process. One of the trends that is in the forefront of small and medium enterprises in cutting operations all over the world is high feed machining [4-5].

High Feed Machining (HFM) is a method incorporated into progressive machining methods. It is three times faster than a conventional machining method. This technology has several advantages and disadvantages which are presented in Table 1. Using a suitable tool with this method, small depths of cut are allowed in conjunction with a high feed per tooth. The result of this machining process is a large amount of material removal. During high feed machining, the depth of cut is small, but the width of cut is optional with respect to the geometry of the tool [6].

The study of high feed machining is currently being dealt with by several experts around the world. For example, in the work of Ji *et al.* [8], the authors present the experimental findings of milling of titanium alloy TC11 using polycrystalline diamond (PCD) cutting tool at high feed rate. Authors Petru *et al.* [9] provide the evaluation of microstructure and microhardness of surface layer after high feed milling in their research. Mihail [10] publishes the research about the manner in which is running the high feed milling process in an orthogonal path pocket milling. The experimental modality follows the robust engineering approach, by the Taguchi Method. In 2015, authors Mwinuka and Mgwatu [11] present study about the tool selection for rough and finish CNC milling operations based on tool-path generation and machining optimisation. In 2015, author Choi [12] describes the influence of feed rate on the fatigue performance. The results demonstrate that a higher feed rate induces more compressive residual stresses and the effect of feed rate increases significantly if the loading is reduced. In 2016, Virginija Gyliene and Valdas Eidukynas present the research study of cutting forces assuming the geometry of the milling cutter in the field of high feed face milling [13]. In 2017, authors Zauskova *et al.* [14] describe the method of triaxial measurement of residual stress after machining the surface of sample by high feed milling technology. They compare the various methods of residual stress analysis after high feed milling. There are not so much studies about the high feed machining of aluminum alloys. In this article there is described the experiments for provision of efficient machining of aluminum alloys with high feed rate.

Table 1 Advantages and disadvantages of HFM technology

Advantages	Disadvantages
• Without high revolution of spindle requirements • Axial direction of cutting forces to the spindle • Reduction of vibration • Increased lifetime of tool and superior cut • It is possible to achieve up to 10 times the feed rate compared to the conventional machining method • Achieving of clean shapes that do not require semi - machining	• It is not possible to apply this method in older machines • Increased risk of vibration commencement • Increased noise level and formation of noise tone component during the cutting process • Fixed clamping of workpiece

2. Materials and methods

One of the appropriate ways to assess the production process is to evaluate its effectiveness. In general, the efficiency of machining is most easily defined by mathematical formulations [15] resulting in quantitative indicators that can subsequently be verified or optimized. The overall efficiency of machining by milling technology was determined by the basic time and cost-effectiveness of machining. Total time efficiency of machining is mathematically formulated as follows:

$$t_{total} = t_{load} + \frac{t_{mach}}{f_{mach}} + \frac{t_{ct}}{N} \tag{1}$$

where t_{total} is total machining time (s), t_{load} is the time taken to load and unload the part to and from machine tool (s), t_{mach} is actual machining time (s), f_{mach} is fraction of the time spent in removing material (s), t_{ct} is tool change time (s), and N is number of parts (-).

$$t_{mach} = \frac{V_{vol}}{fdV} \tag{2}$$

$$N = \frac{fdC^{1/n}}{V_{vol}V^{(1-n)/n}} \tag{3}$$

From this basic mathematical formulation of general time efficiency, the time efficiency for milling was derived in the following form:

$$t_{total} = t_{load} + \frac{1}{f_{mach}}\frac{V_{vol}}{\alpha n_c fdV} + \frac{V_{vol}V^{(1-n)/n}}{n_c fdC^{1/n}}t_{ct} \tag{4}$$

where α is rake angle (°), n_c is number of cutting edges (-), f_d is area of cut (mm²), f_{dV} is volume removal rate (mm³), V_{vol} is volume of material to be removed by milling [mm³], C is Taylor coefficient (-), and V is linear cutting speed (m/s).

For the overall economic efficiency of machining in general, the following mathematical formula applies:

$$C_p = (M_t + M_w)t_{total} + \frac{V_{vol}V^{(1-n)/n}}{fdC^{1/n}}C_t \tag{5}$$

where C_p is the cost of manufacture (€), M_t is the charge rate (€), M_w is the labour charge rate (€), and C_t is the cost of consuming cutting edges (€).

From this basic mathematical formulation of general economic efficiency, the economic efficiency for milling was subsequently derived in the following form:

$$C_p = (M_t + M_w)t_{total} + \frac{V_{vol}V^{(1-n)/n}}{n_c fdC^{1/n}}C_t \tag{6}$$

$$M_t = \frac{C_i}{120000 f_0 n_s}\left[\frac{1}{Y} + (f_i + f_m\right] \tag{7}$$

where C_i is the initial purchase price (€), f_0 is fraction of n_s 8-hour shifts a day (n = 1, 2 or 3), 250 day in year, Y is number of years (-), f_i is fraction of the purchase price, and f_m is typically rises as the inflation rate of an economy increases.

High strength aluminum alloy (EN AW-AlZn5.5MgCu) was used to carry out the experiments. This alloy is composed of aluminum, zinc, magnesium and copper, the zinc being the alloying element. This alloy has excellent machinability; its surface is hard and is therefore suitable for etching of structures. It is also characterized by excellent surface polishability. The high strength aluminum alloy is mainly used for production of foam moulds, blow moulds, machine parts, press and cutting tools. The mass fraction of the alloying elements of the alloy is shown in the Table 2.

Five types of milling tools were used in the experiments, the characteristics of which are given in the Table 3. The pictures of individual tools are presented in the Fig. 1.

Table 2 The mass fraction of the alloying elements of the alloy EN AW ALZn5.5MgCu

Alloy	Alloying elements – mass fraction (%)								
	Fe	Si	Cu	Mg	Mn	Cr	Ti	Zn	other
EN AW ALZn5,5MgCu T651	0.50	0.40	1.20 2.00	2.10 2.90	0.30	0.18 0.28	0.20	5.10 6.10	0.15

Table 3 Technical specification of using tools

Tools	Diameter (mm)	No. of teeth (mm)	Max. cutting depth (mm)
Milling head	40	4	9
Face cutter	40	4	9
Plunge-cutting Φ42	42	3	15
Plunge-cutting Φ32	32	3	15
High feed milling cutter	20	4	1.15
Monolith cutter	12	4	1.5

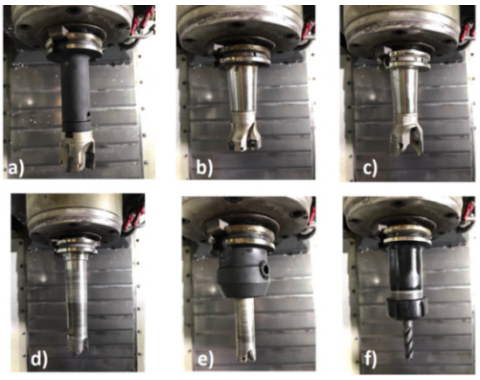


Fig. 1 Used tools: (a) Milling head, (b) Face cutter, (c) Plunge-cutting Φ42, (d) Plunge-cutting Φ32, (e) High feed milling cutter, (f) Monolith cutter

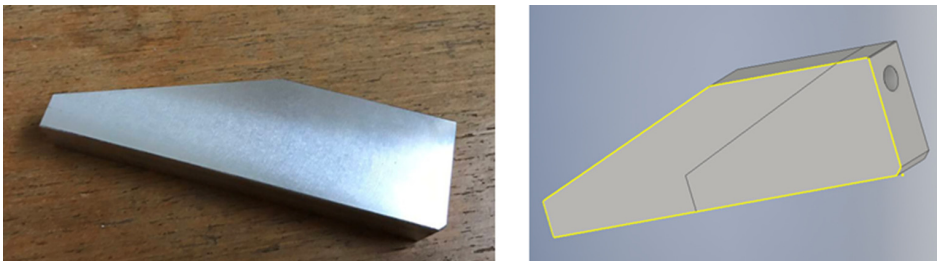


Fig. 2 Produced component (left) and built 3D model (right) – EN AW-AlZn5.5MgCu

During the experiments aimed at analysing the efficiency of high feed machining of aluminum alloys, the simulation method was used. Through this method and the 3D CAD program a basic machining program was created. This program was subsequently imported into the CNC machining centre, which was used for machining of a 115 × 35 × 20 mm blank. The display of the produced component is shown in the Fig. 2.

3. Description of the experiments conducted

To ensure the reliability of the results achieved, five basic experiments were carried out:

- The experiment No. 1– using the face milling cutter
- The experiment No. 2 – using the plunge – cutting cutter Φ42
- The experiment No. 3 – using the plunge – cutting cutter Φ32
- The experiment No. 4 – using the high feed milling cutter
- The experiment No. 5 – machining by conventional milling technology

In all experiments, in addition to the above, a milling head and a monolith cutter were used. In the first four experiments, it is about a progressive high feed milling technology. The experiment No. 5 is about machining with conventional milling technology, which serves for subsequent comparison of high feed milling with conventional milling. As the high feed milling is performed in the first four experiments, the feed rate values should be above 2,000 mm/min and the depths of cut should not exceed 2 mm. An example of clamping of the blank during the experiment is illustrated in the Fig. 3.

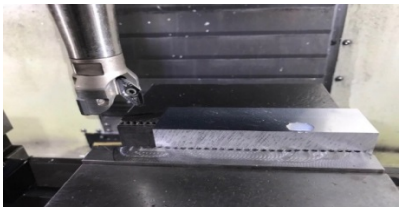


Fig. 3 Experimental realization – clamping of the blank

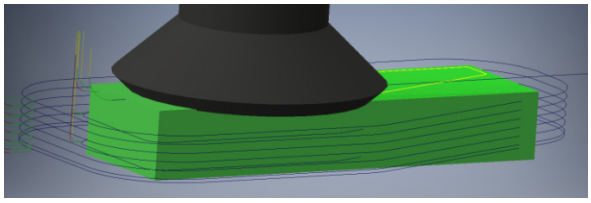


Fig. 4 Simulation of the head alignment of the produced component

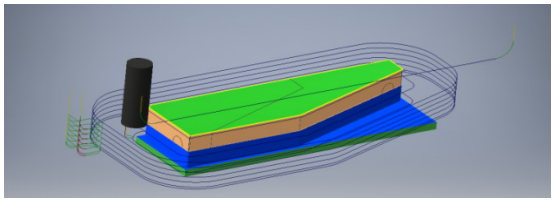


Fig. 5 Simulation of contour milling of the produced component

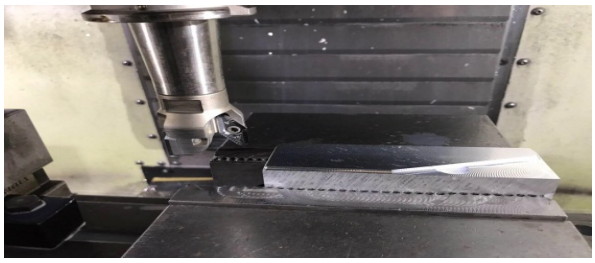


Fig. 6 Experimental realization – contour milling

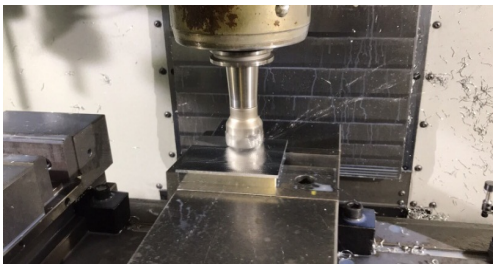


Fig. 7 Experimental realization (second clamping) – head machining

The component was produced in two clampings. During the first clamping, the head was aligned and the simulation is captured in the Fig. 4. The head was aligned using the 40 mm diameter milling head with 4 teeth. Alignment of the head with a total thickness of material removal of 0.5 mm was performed during all five experiments.

After aligning the head, the contour milling, i.e. the circumference of the produced component (Fig. 5), was realized. Contour milling was performed in accordance with a pre-assembled simulation of the tool movement and the required thickness of the material removal (Fig. 6). In this operation, milling cutters of different types and diameters were used according to the specific order of the experiment.

In Fig. 7, the second clamping is different for the processes of the progressive and conventional milling method. In the processes in which the high feed milling method (experiments 1-4) is used, the second clamping is composed of two operations – roughing of the head and deburring to the required thickness and surface quality. In conventional milling, only one operation is performed in the second clamping – roughing of the head to the desired dimension with a gradual removal of 1 mm thick material.

During the experiments, individual cutting conditions were calculated for specific tools, the chosen machine and material. The calculated cutting conditions can be divided into three basic areas. The cutting conditions, that were set for all experiments in the head roughing and material removal operations to achieve the required quality are the first area. The particular values of the cutting parameters and the identification of the tools used are given in the Table 4.

Cutting parameters in the second area are determined individually for each experiment in contour roughing operations (experiments No. 1, No. 2, No. 3, No. 4). The individual values of the cutting parameters are given in Table 5.

Table 4 Cutting conditions for selected operations in all experiments

Operation / Parameter	Face roughing	Material reduction for achieving the required quality
Milling cutter	Milling head	Monolith cutter
Diameter (mm)	40	12
No. of teeth	4	4
Cutting speed (m/min)	180	120
Travel speed (mm/min)	480	768
Cutting depth (mm)	1.8	1.5
Speed (min^{-1})	1500	3200
Feed per tooth (mm)	0.08	0.06

Table 5 Cutting conditions for contour roughing: Experiment No. 1, No. 2, No. 3, No. 4

	Experiment No. 1	Experiment No. 2	Experiment No. 3	Experiment No. 4
Diameter (mm)	40	42	32	42
No. of teeth	4	3	3	3
Cutting speed (m/min)	200	550	500	550
Travel speed (mm/min)	3200	10600	2160	10600
Cutting depth (mm)	1	2	2	2
Speed (min^{-1})	4000	4100	6000	4100
Feed per tooth (mm)	0.2	0.85	0.12	0.85
Total depth of material reduction (mm)	15	15	15	15

The third area determines the conditions and cutting parameters for the experiment No. 5 – machining by conventional milling technology. When machining by conventional milling technology, other conditions were used than in previous experiments. The individual conditions in specific operations are found in the Table 6 and Table 7. The cutting conditions for the first clamping are shown in the Table 6 and the cutting conditions for the second clamping are given in the Table 7.

Table 6 Cutting conditions, experiment No. 5 – the first clamping

	Face alignment	2×contour roughing	Material reduction for achieving the required quality
Milling cutter	Milling head	Monolith cutter	Monolith cutter
Diameter (mm)	40	12	12
No. of teeth	4	4	4
Cutting speed (m/min)	120	120	120
Travel speed (mm/min)	230	768	768
Cutting depth (mm)	1.5	1.5	1.5
Speed (min^{-1})	1000	3200	3200
Feed per tooth (mm)	0.08	0.06	0.06

Table 7 Cutting conditions, experiment No. 5 – the second clamping

Operation	Roughing of face milling
Milling cutter	Milling head
Diameter (mm)	40
No. of teeth	4
Cutting speed (m/min)	120
Travel speed (mm/min)	230
Cutting depth (mm)	1
Speed (min^{-1})	1000
Feed per tooth (mm)	0.08

4. Results and discussion

The experiments performed were evaluated in terms of overall time and economic efficiency. To evaluate overall efficiency, the material used, machining technology and overall manufacturing system management must be taken into account. Evaluation of the overall efficiency is based mainly on Eqs. 4 and 6.

In order to meet the stated objective which is assessing the overall time efficiency of aluminum alloy machining, the basic machining time for one piece of manufactured component was determined in individual experiments, as shown in the Table 8 and Fig. 8.

Table 8 Machining time for one piece of manufactured component

No. of experiment	Time τ_s (min)
1	3.47
2	1.51
3	1.54
4	4.24
5	10.71

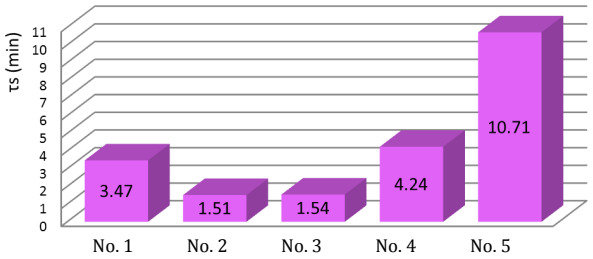


Fig. 8 Graphical interpretation of machining time for one component

Table 9 Total production time of experiments for one piece of manufactured component

No. of Experiment	Total production time (min)
1	4.97
2	3.01
3	3.04
4	5.74
5	12.21

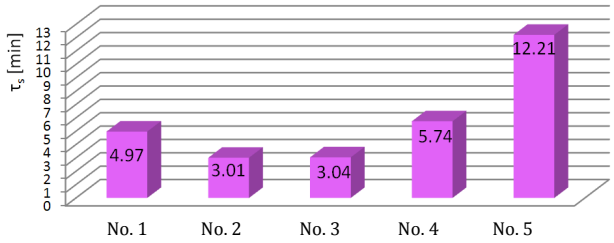


Fig. 9 Graphical interpretation of total production time of individual experiments in the production of one component

From the results of the machining time for one component it can be stated that the best choice from the viewpoint of minimizing the net production time is the second experiment using high-feed milling during which the component was produced (net machining time) in 1.51 min. The worst machining time was achieved by conventional milling method (experiment No. 5).

The overall assessment of the time effectiveness of individual experiments consists of an assessment of the total production time of one component presented in the Table 9 and Fig. 9 and batch production. Comparison of the machining results obtained using high feed milling and conventional milling is also included in the overall assessment of the time-effectiveness of the individual experiments.

As seen in the previous chart and summary table, the most appropriate experiment in the production of one component is the experiment No. 2 – high speed milling, in which a 42 mm plunge-cutting router was used at cutting speed of 550 m/min.

For a more adequate assessment of the manufacturing process, the production time for a single batch production of 1 200 pieces was set. The results of the production time are interpreted in the Table 10 and Fig. 10.

The graphical interpretation of total production time for the batch production in individual experiments confirmed the results achieved for one piece of the manufactured component. As seen in the assessment of the production time of the batch production that is equal to 1200 pieces, the best use of the tools and cutting conditions is described in the experiment No. 2.

Table 10 Total production time of experiments – production batch

No. of experiment	Production time – one batch (min)
1	5964
2	3612
3	3648
4	6888
5	14652

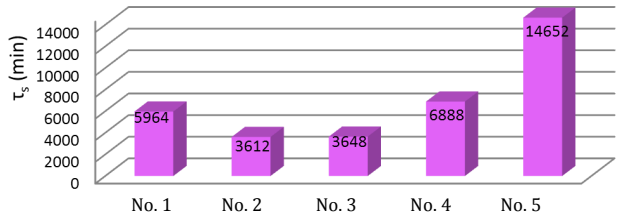


Fig. 10 Graphical interpretation of total production time for experiments – production batch

In order to achieve complete results in the study of the time efficiency of machining of aluminum alloys, an evaluation of conventional milling with realized experiments No. 1-4, in which the high feed milling method was applied, was subsequently implemented. The assessment is carried out by determining the difference among the individual experiments of high speed and conventional milling, as well as the percentage.

The evaluation presented in Table 11 shows that each of the experiments using high feed milling technology is at least two times more time efficient than experiments using the conventional milling method of the component.

Table 11 Comparison of experiment results conducted by HFC method and conventional milling

	HFC method – Experiment No. 1	HFC method – Experiment No. 2	HFC method – Experiment No. 3	HFC method – Experiment No. 4
τs (min) – one piece – HFC methods	4.97	3.01	3.04	5.74
τs (min) – one piece – Conventional milling (Experiment No. 5)	12.21	12.21	12.21	12.21
Time difference	7.24	9.2	9.17	6.47
Percentage	59.296	75.348	75.102	52.989

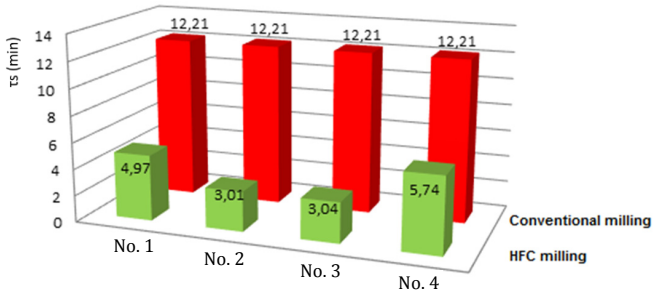


Fig. 11 Graphical comparison of experiments conducted by HFC method and conventional milling

As shown in Fig. 11, after an overall assessment of the time efficiency, it can be stated that of all experiments considered, in order to produce the product in the shortest possible time, it is best to choose high feed milling – namely the parameters specified in the experiment No. 2 which is 75.348 % more time efficient than conventional milling of the material.

The parts of Eqs. 6 and 7 were used to assess the economic efficiency of high feed machining of aluminum alloys. Determination of investment costs for roughing operations is presented in the Table 12.

The price of plates when using a 42 mm plunge-cutting router, which was used in the most time efficient experiment (experiment No. 2), is because the cutting inserts used are made of the material that is the most suitable for machining of aluminum and its alloys. This tool in combination with the given cutting inserts has the longest lifetime for the given parameters, approximately 6000 hours.

To ensure the economic process, it is necessary to choose the tool with the highest life expectancy to delay the additional investment costs of purchasing the tools. Due to the fact that the lifetime of the other tools used is around 3000 hours, the use of alternative No. 2 is considered to be the most appropriate.

Table 12 Determination of investment costs for roughing operations

	No. of plate	Price of tool (€)	Unit price of plate (€)	Price for all plates (€)	Total price of tool with plates (€)
Face milling Φ40	4	380	7	28	408
Plunge-cutting Φ42	3	400	12	36	436
Plunge-cutting Φ32	3	360	8	24	384
High feed milling cutter Φ20	4	160	6.5	26	186

Table 13 Determination the number of pieces produced per hour

No. of experiment	Total production time – one piece of component (min)	No. of pieces produced per hour
1	4.97	12.07
2	3.01	19.93
3	3.04	19.74
4	5.74	10.45
5	12.21	4.91

The value of the produced component is € 2.5 without the material used. Given the fact that in practice the overhead costs are € 20, it is necessary to produce at least 8 pieces in 1 hour to guarantee the efficiency and profitability of the production process. On the basis of the previous determination of time efficiency, it is possible to determine the number of pieces produced per hour in individual experiments. This determination is presented in the Table 13.

According to the criteria set and financial options, it was necessary to produce at least eight pieces of components in one hour to ensure the profitability of the production process. According to the above calculations, the second alternative (the HFC machining method) was confirmed as the most appropriate option in terms of economic efficiency of production.

The graph in the Fig. 12 shows the results obtained from the previous table, with the red line marking a minimum requirement for the number of pieces to be manufactured to ensure the profitability and economic efficiency of the production process.

Based on the evaluation of the economic efficiency of individual experiments, the parameters used in the experiment No. 2 can be considered as the most appropriate solution. This proposal is the best suited both economically and in terms of time efficiency.

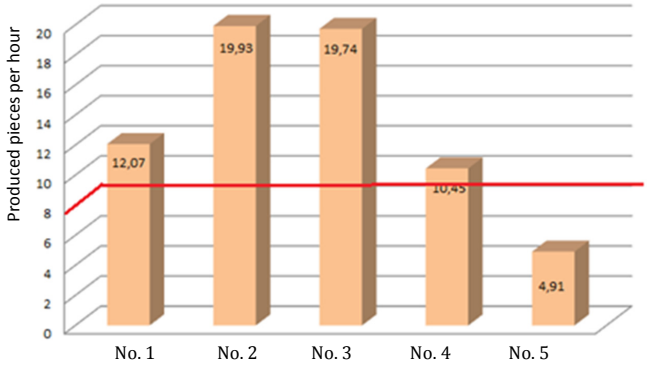


Fig. 12 Graphical interpretation of produced pieces per hour

5. Conclusion

Aluminium alloys are among the primary materials used in the transport industry, mainly in the automotive and aircraft industry [16-18]. Applied research oriented to solving partial problems in manufacturing processes of mass production to which the transport industry undoubtedly belongs, it is currently dealing with time efficiency of production [19-21]. Reducing production time has a direct impact on the total cost of production [22-23]. The high feed machining method undoubtedly ranks into the progressive machining technologies, the use of which the aforementioned factors can be achieved. They were observed the key factors such as time efficiency, reduction of production time, total cost of production, cost reduction, economic efficiency and

cost-effectiveness. The results obtained show that the production process is most efficient when using a 42 mm plunge-cutting router in the contour roughing phase in terms of both time and economy. Using this milling cutter, it is possible to produce 19 pieces in one hour, which is more than half of the number required. The production of components under the conditions using this type of high-feed milling cutter is more than 75 % shorter than production by a conventional method. In conclusion, although the machining of aluminum and aluminum alloy materials by means of high feed milling technology is more costly in terms of investment costs, this slightly negative property is sufficiently balanced in terms of time efficiency and productivity. The research carried out in the presented article points out that the machining by the progressive high feed milling method is a suitable option for smaller companies that need to achieve a faster production of components of the required quality at an affordable price.

Acknowledgement

Presented article was supported by research grant VEGA 1/0492/16 (Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic).

References

[1] Čepová, L., Šoková, D., Malotová, Š., Gapinski, B., Čep, R. (2016). Evaluation of cutting forces and surface roughness after machining of selected materials, *Manufacturing Technology*, Vol. 16, No. 1, 45-48.

[2] Mitaľová, Z., Mitaľ, D., Botko, F. (2016). Measuring of roughness and roundness parameters after turning of composite material with natural reinforcement, In: *Science report: Project CIII - PL-0007: Research on modern systems for manufacture and measurement of components of machines and devices,* Kielce: Wydawnictwo Politechniki Świętokrzyskiej, Poland, 49-58.

[3] Iwaszko, J., Kudla, K., Fila, K., Caban, R. (2017). Application of FSP technology in formation process of composite microstructure in AlZn5.5MgCu aluminium alloy surface layer reinforced with SiC particles, *Composites Theory and Practice*, Vol.17, No. 1, 51-56.

[4] Wu, H., Zhang, S. (2015). Effects of cutting conditions on the milling process of titanium alloy Ti6Al4V, *The International Journal of Advanced Manufacturing Technology*, Vol. 77, No. 9-12, 2235-2240, doi: 10.1007/s00170-014-6645-2.

[5] Napiorkowski, J., Mikolajczak, P., Legutko, S., Krolczyk, J. (2017). Developing of wear model of construction materials in abrasive soil pulp employing discriminant analysis, *Tehnički Vjesnik – Technical Gazette*, Vol. 24, Supplement 1, 15-20, doi: 10.17559/TV-20140422230704.

[6] Davim, J.P. (2011). *Modern machining technology: A practical guide*, Elsevier Science & Technology, Cambridge, UK.

[7] Lei, F., Xu, Q., Zhang, G. (eds.) (2017). *Machinery, materials science and engineering applications: Proceedings of the 6th international conference on machinery, materials science and engineering applications (MMSE 2016),* Wuhan, China, CRC Press/Balkema.

[8] Ji, W., Liu, X., Wang, L., Sun, S. (2015). Experimental evaluation of polycrystalline diamond (PCD) tool geometries at high feed rate in milling of titanium alloy TC11, *The International Journal of Advanced Manufacturing Technology*, Vol. 77, No. 9-12, 1549-1555, doi: 10.1007/s00170-014-6517-9.

[9] Petru, J., Cep, R., Grepl, M., Petrkovska, L. (2011). Effect of high feed milling on the microstructure and microhardness of surface layer, *Annals of DAAAM for 2011 & Proceedings of the 22nd International DAAAM Symposium*, Vol. 22, No. 1, 999-1000.

[10] Mihail, L.A. (2010). Dynamic mill's deflection for high feed machining on orthogonal directions, In: *Proceedings of the 9th WSEAS international conference on Signal processing, robotics and automation,* Cambridge, UK, 69-73.

[11] Mwinuka, T.E., Mgwatu, M.I. (2015). Tool selection for rough and finish CNC milling operations based on toolpath generation and machining optimisation, *Advances in Production Engineering & Management,* Vol. 10, No. 1, 18-26, doi: 10.14743/apem2015.1.189.

[12] Choi, Y. (2015). Influence of feed rate on surface integrity and fatigue performance of machined surfaces, *International Journal of Fatigue,* Vol. 78, 46-52, doi: 10.1016/j.ijfatigue.2015.03.028.

[13] Gylienė, V., Eidukynas, V. (2016). The numerical analysis of cutting forces in high feed face milling, assuming the milling tool geometry, *Procedia CIRP*, Vol. 46, 436-439, doi: 10.1016/j.procir.2016.03.132.

[14] Zauskova, L., Czan, A., Sajgalik, M., Drbul, M., Rysava, Z. (2017). Triaxial measurement of residual stress after high feed milling using X-ray diffraction, *Procedia Engineering*, Vol. 192, 982-987, doi: 10.1016/j.proeng.2017.06.169.

[15] Childs, T., Maekawa, K., Obikawa, T., Yamane, Y. (2000). *Metal machining: Theory and applications*, John Wiley & Sons, New York, USA, doi: 10.1016/C2009-0-23990-0.

[16] Valíček, J., Harničárová, M., Öchsner, A., Hutyrová, Z., Kušnerová, M., Tozan, H., Michenka, V., Šepelák, V., Mitaľ, D., Zajac, J. (2017). Quantifying the mechanical properties of materials and the process of elastic-plastic deformation under external stress on material, *Materials,* Vol. 8, No. 11, 7401-7422, doi: 10.3390/ma8115385.

[17] Knežo, D., Andrejiová, M., Kimáková, Z., Radchenko, S. (2016). Determining of the optimal device lifetime using mathematical renewal models, *TEM Journal,* Vol. 5, No. 2, 121-125.
[18] Knapčíková, L. (2013). Examination of surface of composite materials by atomic force microscopy, *Strojárstvo*, Vol. 17, No. 12, 58-59.
[19] Lehocká, D., Hlavatý, I., Hloch, S. (2016). Rationalization of material flow in production of semitrailer frame for automotive industry, *Tehnički Vjesnik – Technical Gazette*, Vol. 23, No. 4, 1215-1220, doi: 10.17559/TV20131113100109.
[20] Karpus', V.E., Ivanov, V.A. (2008). Universal-composite adjustable machine-tool attachments, *Russian Engineering Research*, Vol. 28, No. 11, 1077-1083, doi: 10.3103/S1068798X08110105.
[21] Katahira, K., Matsumoto, Y., Komotori, J., Yamazaki, K. (2017). Experimental investigation of machinability and surface quality of sapphire machined with polycrystalline diamond micro-milling tool, *The International Journal of Advanced Manufacturing Technology*, Vol. 93, No. 9-12, 4389-4398, doi: 10.1007/s00170-017-0881-1.
[22] Masood, I., Jahanzaib, M., Haider, A. (2016). Tool wear and cost evaluation of face milling grade 5 titanium alloy for sustainable machining, *Advances in Production Engineering & Management*, Vol. 11, No. 3, 239-250, doi: 10.14743/apem2016.3.224.
[23] Neacşu, M.I., Chiriac, R.E., Chiriac, A., Pandia, O., Saracin, I. (2017). Experimental research on the influence of soaking aging type on some mechanical properties of the alloy AlZn5,7MgCu, *Metalurgija*, Vol. 56, No. 1-2, 215-218.

APEM journal
Advances in Production Engineering & Management
Volume 13 | Number 4 | December 2018 | pp 466–478
https://doi.org/10.14743/apem2018.4.304
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Hybrid fruit fly optimization algorithm for solving multi-compartment vehicle routing problem in intelligent logistics

Wang, C.L.[a], **Li, S.W.**[a,*]

[a]School of Business, Shandong Normal University, Ji'nan, P.R. China

ABSTRACT

The purpose of this study was to tackle multi-compartment vehicle routing problem in intelligent logistics with the fruit fly optimization algorithm (FOA). A hybrid FOA (HFOA) integrated with three local search methods (2-opt, swap and insert) was adopted to solve the multi-compartment vehicle routing problem (MCVRP) in intelligent logistics by applying discrete space optimization problems. The numerical experiments show that the HFOA algorithm has improved the performance for all proposed problems, including improving the total path length and enhancing the solution quality. The improvement rate in total path length shifts from 3.21 % at 50 customers to 9.83 % at 150 customers indicating that this HFOA is more effective in largescale. The HFOA integrated with 2-opt, swap and insertion elevates the solution quality from 11.86 % to 17.16 % displaying the advantages. The effectiveness and stability of the proposed algorithm shed new light on the routing of MCV distribution problems in intelligent logistics.

ARTICLE INFO

Keywords:
Intelligent logistics;
Vehicle routing problem (VRP);
Multi-compartment vehicle (MCV);
Bionic optimization;
Fruit fly optimization algorithm (FOA)

**Corresponding author:*
3472181435@qq.com
(Li, S.W.)

Article history:
Received 3 September 2018
Revised 4 December 2018
Accepted 7 December 2018

1. Introduction

Intelligent logistics is a modern comprehensive logistics system supported by information technology. It realizes system perception, comprehensive analysis, timely processing and self-adjustment in transportation, warehousing, packaging, handling, circulation processing, distribution, information service and so on. In the modern logistics industry, the logistics distribution carries two main features: the various and numerous products to be transported, and the rapidly increasingly transport demand. For a distribution system, the total cost relies heavily on the path length of the vehicle. In light of these, vehicle routing must be optimized to enhance the transport efficiency and lower the distribution cost. This contributes to the focus on the vehicle routing problem (VRP) in the intelligent logistics industry and the academia.

Since 1959, the VRP has always been a hot topic in operations research (Dantzig and Ramster) [1]. Over the years, computer simulation, logistics planning and many other strategies have been introduced to the VRP research, yielding fruitful results. Recent years has seen the rise of the multi-compartment VRP (MCVRP) with capacity constraints. This problem mainly considers the multiple compartments of the vehicle, and ensures the separation between different products.

Both the VRP and the MCVRP are NP-hard problems, for the traveling salesman problem (TSP), a special case of the VRP, has been proved as NP-hard (Garey, 1979) [2]. It is difficult to solve a large-scale NP-hard problem with traditional mathematical optimization algorithm.

Hence, more and more attention has been paid to bionic optimization, especially the fruit fly optimization algorithm (FOA), aiming to find a good solution to the VRP in a short time.

Unlike traditional optimization algorithm, the bionic optimization, such as the FOA, is an emerging evolutionary algorithm, which is easy to understand, operate and implement. It has been applied to many continuous space optimization problems, namely, finance, electric load forecast, logistics service, and parameter tuning [4-6]. Nevertheless, the FOA, as a global optimization algorithm, is prone to the local optimum trap. The convergence is particularly slow and inaccurate in the later stage of the search. To solve the defect, some scholars have adopted the chaos theory [7] and bacterial chemotaxis algorithm [8] to improve the FOA.

Despite the good effect of the FOA and its improved versions, the existing research emphasizes continuous space optimization problems over discrete space optimization problems like the combinatorial optimization problem. To make up for the gap, this paper attempts to solve the discrete vehicle routing problem by a hybrid FOA (HFOA) based on the MCV and local search.

The remainder of this paper is organized as follows: Section 2 reviews the previous research on the VRP, and presents some solutions to the MCVRP; Section 3 details the model of the MCVRP; Section 4 presents the HFOA based on the local search; Section 5 verifies the effect of the HFOA through computational experiments; Section 6 wraps up this paper with some meaningful conclusions.

2. Literature review

The vehicle routing problem is a type of optimal scheduling problem that studies how to optimize the transportation cost by rationally planning the driving route. Its related theories and algorithms have important application value for reducing logistics costs, so it has always been a research hotspot in the field of operations research and combinatorial optimization. Over the years, vehicle routing problems have spawned numerous research branches, such as open vehicle routing problems, multi-site vehicle routing problems, loading and unloading vehicle routing problems, vehicle routing problems with time windows, and periodic vehicle routing problems. And have obtained a lot of research results, At the same time, the vehicle routing problem is also widely used in all aspects of production and life, such as letter delivery, cargo distribution, vehicle scheduling, etc., which has produced huge economic benefits.

As a generalized form of the TSP, the VRP (Dantzig and Ramser, 1959) aims to transport products from the distribution centre (warehouse) to customers with a fleet of vehicles, and fulfil customer demands at the minimal cost. In the course of application, the VRP with capacity constraints (CVRP) emerges due to the limited transport capacity of vehicles. To minimize the total distance covered by all vehicles, the CVRP assumes that each vehicle has a constant capacity Q to serve a fixed number of customers, each of which has a fixed demand, that each vehicle accesses each customer only once, and that the total demand on any path does not surpass the vehicle capacity.

Considering the varied transport requirements of customers, some scholars have presented the MCVRP under the following premises: each customer has a fixed and known demand for each product, different products are stored in separate compartments of the same vehicle during transport, each vehicle has a fixed number of compartments, and each compartment has a certain capacity limit. In addition, the total customer demand for any product should not exceed the capacity of the vehicle that carries the product, when a customer is assigned to a distribution path. The goal of the MCVRP is also to find the minimal total distance of distribution.

The multi-compartment configuration is essential to transporting various products that cannot be mixed together all at once. For example, the MCVRP has been applied to the transport of food. The refrigerated foods and non-refrigerated foods are stored in different compartments of the same vehicle. Chajakis and Guignard (2003) explored two integer programming models of two-compartment vehicles, and discussed the decision-making of assigning the customer to the distribution path [9]. The MCVRP has also been applied to the transport of fuels, such that different types of fuels are stored in tanks of varied capacities. For instance, Avella *et al.* (2004) developed a branch and bound algorithm based on set partitioning [10]. Fallahi *et al.* (2008) solved multi-tank transport [11] with memetic algorithm and tabu search algorithm.

For better performance, the MCVRP has been frequently enhanced by local search and heuristic algorithm. Focusing on the CVRP, Chen *et al.* (2009) proposed an iterative local search algorithm based on multiple neighbourhoods [12]. Muyldermans and Pang (2010) compared the MCV with single-compartment vehicle (SCV), revealing the advantage of the former in simultaneous transport of different types of garbage separately from different locations to the collection centre. Specifically, they solved the classified garbage transport problem by such local search methods as 2-opt, crossover, swap and path redistribution, and contrasted the results with those of Fallahi [13]. Avella *et al.* (2004) assumed that the demand of each customer for certain products is inseparable, and that multiple vehicles can access the same customer to fulfil the demand of different products. To solve the CVRP in garbage collection network, Reed *et al.* (2014) [14] proposed an improved ant colony (AC) system with 2-opt local search, and solved the MCVRP in which each customer can be accessed only once by a vehicle. Gajpal and Abad (2009) created an AC algorithm to solve the VRP with simultaneous transmission [15]. Balseiro *et al.* (2011) combined the interpolation method with the AC algorithm to solve the VRP with time window constraints [16]. Based on the AC and tabu search, Cruz *et al.* (2013) proposed a sequential algorithm to solve the VRP problem with time window constraints, which transports various products with heterogeneous vehicles [17]. Valiček *et al.* (2017) and Tan *et al.* (2015) designed a heuristic algorithm which combines the AC system and 2-opt to tackle the VRP, and proved the algorithm as effective in solving the VRP and its deformation problem [18-19].

Inspired by Reed *et al.* (2014), the author extended the MCVRP using the multiple compartments of the vehicle. In the new problem, each vehicle transports various types of products in different compartments from the distribution centre (warehouse) to the customer. The constraints are as follows: all vehicles in the fleet are the same, each vehicle accesses to a specific group of customers, and each customer is accessed once only by each vehicle. For the minimal total travel distance, the existing FOA was enhanced by local search to assign customers rationally to a single path, and determine the order of these customers.

3. Materials and methods

3.1 Modelling of multi-compartment vehicle routing problem

In an actual logistics distribution network (Fig. 1), the MCVRP involves the following elements: a single distribution centre (warehouse), a fleet of MCVs, and several customers in need of distribution services. Note that each customer wants various products, the demand of each product is fixed, and each compartment has a fixed capacity. To fulfil the customer demand, each vehicle must leave from the distribution centre, access some customers and return to the centre. Each customer can only be accessed once by a vehicle. Given the distances between network vertices, the goal is to find the path of each vehicle that contributes to the minimum total travel distance. Table 1 lists the symbols associated with the mathematical model of the MCVRP and their meanings.

The undirected graph $G = (V, A, C)$ was employed to formally describe the distribution network of the MCVRP. Let $V = \{i | i = 0,1, \dots, n\}$ be the set of vertices in graph, with $i = 0$ being the distribution centre (warehouse), , in which the zero vertex is distribution $N = \{1, \dots, N\}$ be the set of customers served by k vehicles, and $A = \{(i, j) | i, j \in V\}$ the set of edges linking up the network vertices.

It is assumed that the vehicles are the same and located in the distribution centre (warehouse) at the beginning. Suppose each vehicle has p compartments, and p equals the number of products to be delivered. The number of product p demanded by customer I is denoted as q_{ip}.

For each customer, he/she should be accessed by a vehicle only once; for each vehicle, there should be a set of customers to be accessed in a strict sequence; for a given product, the total demand of the set of customers should not surpass the products to be deliver for customer i capacity Q_P of the compartments; for each path, the maximum length should not exceed L.

Table 1 Symbols associated with the MCVRP model and their meanings

Symbol	Meaning	Symbol	Meaning
V	The vertex of distribution network, including customers and distribution centres	Q_{ip}	Number of product p that customer i need to transport
A	Adjacency matrix of distribution network	Q_p	The carriage capacity of the product p
C	Vertex distance matrix of distribution network	L	Maximum length for any path
N	Set of customers	c_{ij}	Length of $arc(ij)$
K	Set of vehicles	x_{ij}^k	If vehicle k accesses to customer j after visiting customer i, x_{ij}^k=1
p	Set of products, which is equal to the set of compartments in a vehicle	Q_{ip}^k	The total transport capacity of the product p after vehicle k leaves customer i

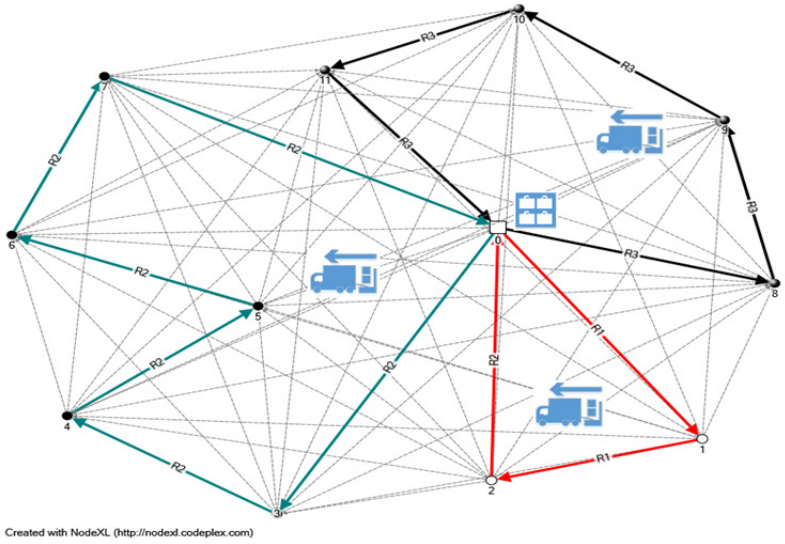


Fig. 1 MCVRP distribution network

Let $C = \{c_{ij}|(i,j) \in A\}$ be the distance matrix of network, where c_{ij} is the distance between customers i and j. The matrix is assumed to be symmetric, that is, $c_{ij} = c_{ji}$ .Let x_{ij}^k be a binary variable, which equals 1 if and only if vehicle k accesses customer j after accessing customer i Let Q_{ip}^k be the total delivery of product p after vehicle k leaves vertex i.

Hence, the MCVRP can be expressed as:

$$Z = \sum_{k\in K}\sum_{i\in V}\sum_{j\in V} c_{ij}\, x_{ij}^k \tag{1}$$

Min

$$\sum_{k\in K}\sum_{j\in N} x_{ij}^k = 1, \forall i \in N \tag{2}$$

$$\sum_{k\in K}\sum_{i\in N} x_{ij}^k = 1, \forall j \in N \tag{3}$$

$$\sum_{i\in N} x_{0i}^k = \sum_{j\in N} x_{j0}^k = 1, \forall k \in K \tag{4}$$

$$q_{ip} \le Q_{ip}^k \le Q_p, \forall i \in N, k \in K, p \in P \tag{5}$$

$$Q_{ip}^{k} - Q_{jp}^{k} + Q_P x_{ij}^{k} \leq Q_p - q_{jp}, \forall i \in N, k \in K, p \in P \tag{6}$$

$$\sum_{i \in V} \sum_{j \in V} c_{ij}\, x_{ij}^{k} \leq L, \forall k \in K \tag{7}$$

$$x_{ij}^{k} \in \{0,1\}, \forall i \in V, j \in V, k \in V, i \neq j \tag{8}$$

Eq. 1 is the objective function that represents the total distance of all vehicles on all paths. Eq. 2 and Eq. 3 ensure that a path should pass through a vertex only once, that is, one vehicle can only access each customer once. Eq. 4 specifies that the path of each vehicle starts and ends at point 0 (warehouse). Eq. 5 and Eq. 6 eliminate the sub-loops and satisfy the capacity and connection demands between two customers. If $x_{ij}^{k} = 0$, Eq. 6 is not required as $Q_{ip}^{k} \leq Q_p$ and $Q_{jp} \leq Q_{jp}^{k}$. If $x_{ij}^{k} = 1$, Eq. 5 and Eq. 6 guarantee that $Q_{ip}^{k} - Q_{jp}^{k} \leq -q_{jp}$, thus eliminating the sub-loops. Eq. 5 also stipulates that the total transport capacity falls within the vehicle capacity after the vehicle accesses vertex i, thus ensuring the capacity demand. Eq. 7 is the constraint of path length. Eq. 8 describes the variable x_{ij}^{k}, which equals 1 if and only if vehicle k accesses customer $j$after accessing customer i.

3.2 Hybrid fruit fly optimization algorithm

The FOA is a popular tool to find the optimal solution to NP-hard problems. It mimics the olfactory and visual functions of fruit flies in the foraging process. By this algorithm, the status of fruit flies is updated by searching the global optimum iteratively. In the course of iteration, once a fruit fly finds a better global optimum, all individuals in the population will gather to its location. The individual update mechanism lowers the population diversity, because the individual information is not shared or inherited, and adds to the risk of local optimum and premature convergence, as the position may not be the global optimum [20-21]. Zheng (2014) proposed a novel hybrid discrete algorithm for permutation flow scheduling problem. In the algorithm, the evolution of each generation consists of 4 phases: olfactory search, visual search, cooperative evolution and annealing [22]. Considering the strengths and weaknesses of the FOA, the author created a HFOA for the MCVRP.

Algorithm principle

Before updating path strength, the HFOA takes account of both the current best individual and the local optimum in the update of individual status. In this way, the algorithm ensures that the fruit flies concentrate towards the global optimum and inherits the local optimum information at the same time. The principle of the HFOA is illustrated in Fig. 2.

The HFOA is implemented in the following steps:

Step 1: Create the initial solution and initialize the path strength.
Step 2: Repeat the following operations until the termination condition.
Step 2.1: Create the path for m fruit flies.
Step 2.2: Perform local search to improve the solutions generated by each fruit fly.
Step 2.3: Update the best solution.
Step 2.4: Update the path strength for all edges against the best solution.
Step 3: Terminate the algorithm and record the best solution.

The HFOA uses fruit flies to construct path solutions to the MCVRP. In each iteration for path construction, each fruit fly performs four basic activities in turns. First, the fruit fly selects the next customer based on the probability function of path attraction, which consists of taste strength and path strength. Second, the fruit fly accesses the *Tabu* list of customers in the current path. Third, the fruit fly updates the residual capacity of the compartment of the vehicle. Fourth, the fruit fly updates the path strength of edges which have been accessed. After that, the local search is introduced to improve the solution quality. Finally, the *Tabu* list is deleted, marking the start of a new iteration. Table 2 shows the symbols of the HFOA and their meanings.

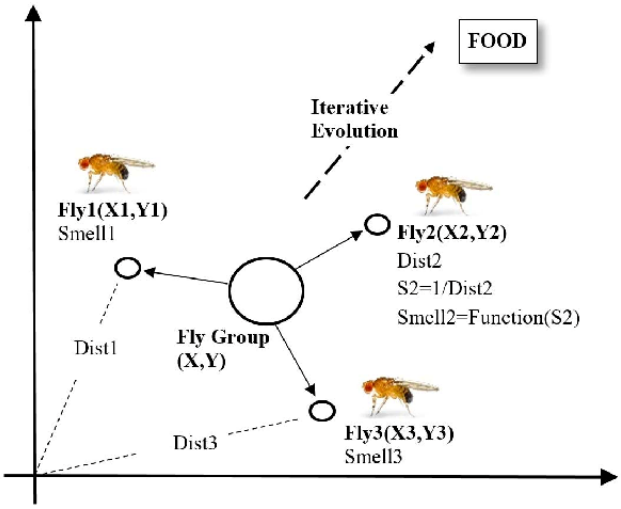


Fig. 2 Principle of the HFOA

Table 2 HFOA symbols and their meanings

Symbol	Meaning	Symbol	Meaning
m	Number of fruit flies	N_i	List of all possible customers which are not accessed by fruit fly
μ_{ij}	The strength of taste, that is, the reciprocal of path ij length	τ_{ij}	The strength of track, that is, the access times of path (ij)
ε_{ij}	The attractive value of path (ij)	P_{ij}	The probability that a customer j will be accessed after the customer i is accessed by a fruit fly.
q	Random variables, to determine whether the customer choice is based on attractive probability or the biggest attraction.	q_0	Threshold of q
L^{best}	The length of the best solution found so far	ρ	Persistence coefficient of path
$Tabu_s$	The tabu list of customers firstly visited by all fruit flies	$Tabu_i$	The tabu list of customers which are accessed by fruit fly i

Initialization

The first step of the HFOA is to generate the initial solution and initialize the strength of each path. The initial solution is generated in two steps. First, the customers are randomly generated according to the topological structure of the network; if the vehicle capacity is enough to meet customer demand, the customers are randomly added to the path of the vehicle; otherwise, the vehicle returns to the distribution centre (warehouse) before accessing the next customer. Fig. 3 is the flow chart of path initialization of the HFOA.

Path creation

Suppose each of the m fruit flies in the HFOA generates a complete path (a complete MCVRP solution). For solution diversify, each fruit fly should select the first customer randomly after leaving the distribution centre (warehouse). Next, each fruit fly should move to the subsequent customers based on the probability of path attraction (i.e. taste strength and path strength).

In the FOA, the taste strength is the reciprocal of the distance between the current location and the origin of the fruit fly. Here, the distance between the current location and the next customer is inversed as the taste strength. It is obvious that the strength is negatively correlated with the distance. Hence, the fruit flies tend to choose customers closer to the origin, laying the basis for minimizing the path length.

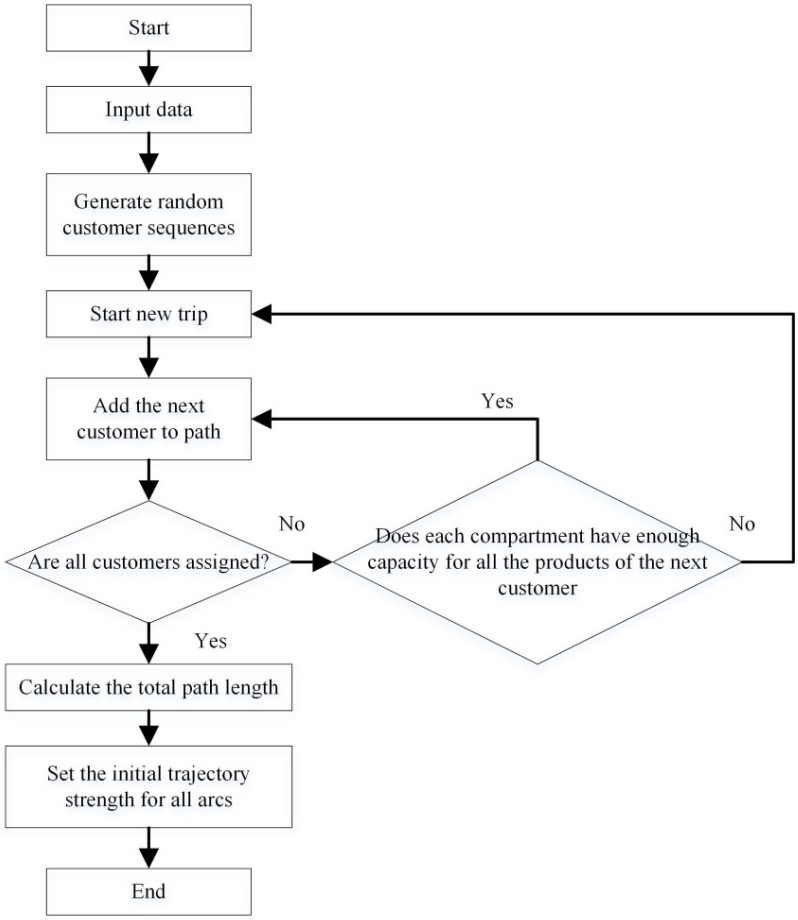


Fig. 3 Flow chart of path initialization in the HFOA

In light of the above, the attraction of edge ij consists of attraction power includes two parts: the path strength τ_{ij}, that is, the frequency of edge ij accessed during the iteration, and the taste strength μ_{ij}, that is, the reciprocal of the length of edge ij . The latter also represents the moving tendency of the fruit fly from i to j. Thus, the path attraction can be expressed as:

$$\varepsilon_{ij} = (\tau_{ij})^{\alpha}(\mu_{ij})^{\beta} \tag{9}$$

where α and β are the indices of path strength and taste strength, respectively. The probability that customer j will be accessed by the fruit fly after customer i is denoted as p_{ij} will be the next customer after the fruit fly visits the customer i:

$$p_{ij} = \begin{cases} \dfrac{\varepsilon_{ij}}{\sum_{l \in N_i} \varepsilon_{il}} & j \in N_i \\ 0 & \text{other} \end{cases} \tag{10}$$

where N_i is the list of all customers not yet accessed by the current vehicle, whose total demand does not exceed the vehicle capacity. According to Eq. 10, the choice probability of the next customer is proportional to the attraction of edge ij.

To accelerate the convergence, a threshold control mechanism was developed for customer selection. Before the fruit fly chooses the next customer j according to Eq. 10, a random variable q is generated such that it is evenly distributed in the interval of [0,1]. The next customer is selected according to Eq. 10 that is evenly distributed on the [0,1] is generated before each fruit fly chooses the next customer j according to the Eq. 10. If $q > q_0$; otherwise, the customer on the most attractive edge is selected. Hereinto, q_0 is a constant threshold indicating whether the next customer is determined according to Eq.10.

Next, each fruit fly continues to add customers in accordance with the above rules until no viable customer is left (N_i is empty). Then, the fruit fly goes back to the distribution centre and starts a new path. If the first customer on the new path is selected according to Eq. 10, then he/she must be the neighbouring vertex to the distribution centre. Thus, the customer search process becomes less diverse.

Local search

Since each trip of a fruit fly is a TSP (Lin, 1965) [23], the travel distance can be shortened by ensuring that the total demand on each path does not exceed the capacity limit. Hence, three local search operations were implemented after creating all the paths, aiming to enhance the solution quality, namely 2-opt, swap and insert:

- 2-opt: The path is divided at two points into three segments. Then, the customer sequence in the middle segment is inverted. After that, all the segments are linked up to reconstruct the path. Suppose that i and j are two non-adjacent customers on a path, and i^+ and j^+ are next adjacent vertices in the path, respectively. 2-opt means the creation of a new path by deleting edges (i, i^+) and $(j, j^+)$and adding edges (i, j) and (i^+, j^+). And (j, j^+), and adding and (i^+, j^+), so we get a new path.
- Swap: The locations of customers i and j in the current path set are switched to generate a new set of paths. Note that i and j may or may not fall onto the same path.
- Insert: Customer i is moved from location p_1 to location p_2 to generate a new set of feasible paths in the current path set. Note that $p_1$and p_2 may or may not fall onto the same path.

Through the above local search operations, the author obtained all feasible solutions and selected the shortest path from them.

Path strength update

In this step, the path strength of all edges is updated based on the optimal solution identified in the local search process. To simulate the actual situation, the path strength is assumed to decay with time.

The path strength is updated in two steps. First, the strength of all edges is reduced to reflect the decay intensity; second, the strength of the best path is increased to drive the fruit flies towards the shortest path. Let ρ be the path persistence coefficient$(0 \leq \rho < 1)$. Then, $1 - \rho$ discloses the path strength decay ratio during the iteration. Thus, the edge ij strength can be updated by the formula below:

$$\tau_{ij}^{new} = \begin{cases} \rho\tau_{ij}^{old} + 1/L^{best} & ij \in best\ route \\ \rho\tau_{ij}^{old} & other \end{cases} \tag{11}$$

where L^{best} is the total length of the best path in each iteration. The product of the path strength and ρ is the residual strength of the path after the decay. Then, the strength of the optimal path is updated by represents the residual strength of the trajectory after the decay, and then the adding $1/L^{best}$ to the residual decay intensity.

Fig. 4 presents the steps of the HFOA for the MCVRP.

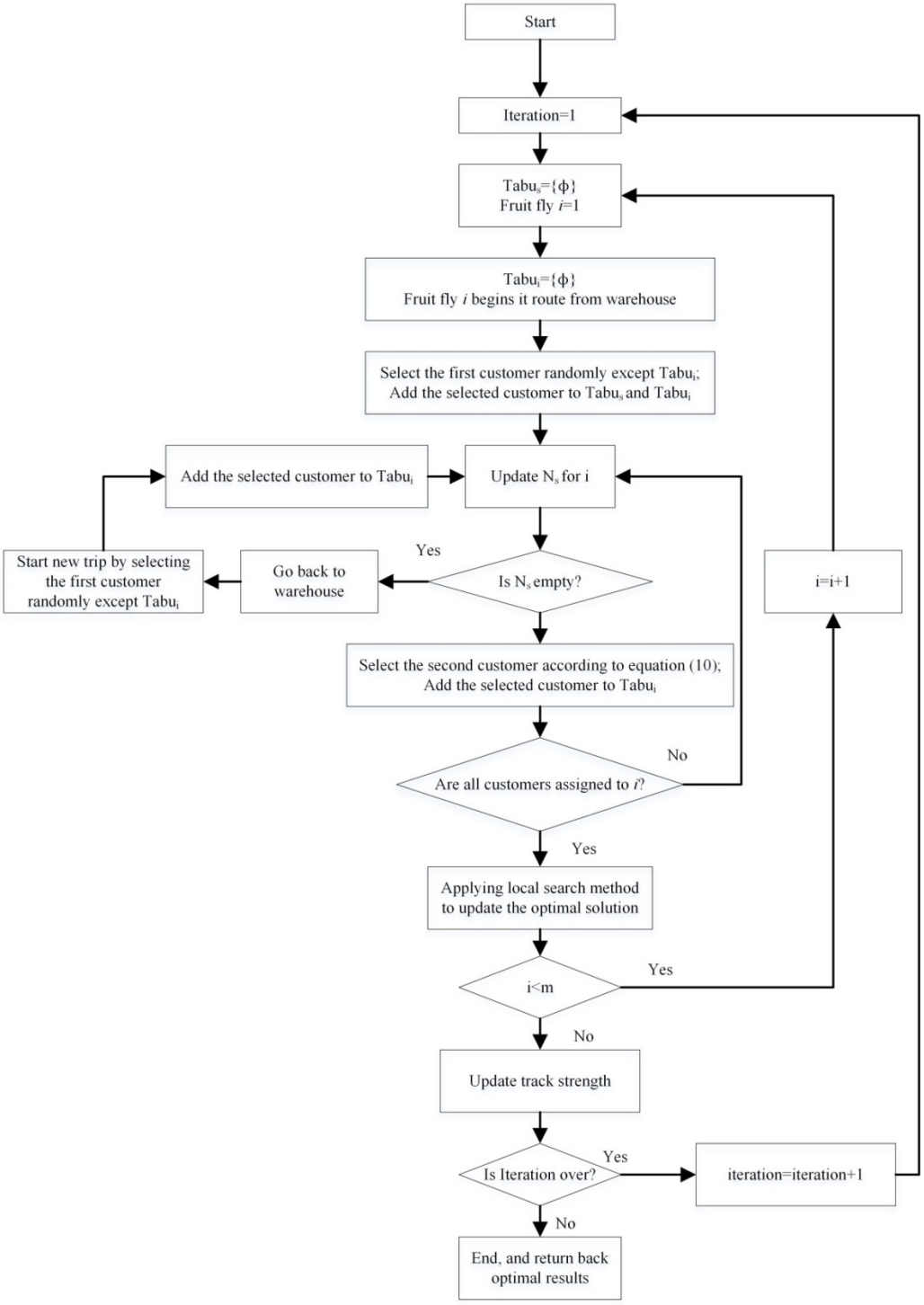


Fig. 4 Flow chart of HFOA for MCVRP

4. Results and discussion

4.1 Data and parameter settings

To validate the HFOA, the benchmark problem proposed by Laporte [24] was investigated, and the proposed algorithm was contrasted with the AC system [25-27] and hybrid ant colony (HAC) algorithm [28]. In view of the capacity constraint, 7 of 14 benchmark problems, denoted as VRPNC1-7, were selected for this research. The customers in VRPNC1-VRPNC5 obey random and even distribution, while those in VRPNC6-VRPNC7 obey aggregated distribution. These problems involve a total of 50-199 customers.

Suppose each vehicle has two compartments, whose capacity ratio is 1: 3. In other words, if the vehicle capacity is Q, then the capacity of the two compartments is $0.25Q$ and $0.75Q$, respectively.

The number of fruit flies in the HFOA, denoted as m, directly bears on both the best path quality and the computing time. After repeated tests, m was set to 20. The other parameters are as follows: Other parameters are: $\alpha = 1, \beta = 2, \rho = 0.9$ and $q_0 = 0.9$ .The experiment contains 100 iterations.

4.2 Effect analysis

In this section, the AC system, the HAC and the HFOA are simulated in Matlab. The simulation results, together with the improvement rate of the total path length, are recorded in Table 3.

As shown in Table 3, the average total length of the AC system, the HAC and the HFOA are respectively 1,053.977 unit length, 986.9929 unit length and 985.6029 unit length. Thus, the HAC improves the average total length by 6.36 % from that of the AC system, while the HFOA improves it by 6.49 % from that level. Thus, the HFOA has a similar effect to the HAC, and outperforms the AC system, indicating the advantage of heuristic algorithms in optimization operation. The slight edge of the HFOA over the HAC is attributed to the strong local search ability. In VRNPC6 and VRNPC7, the three algorithms differ very slightly in the total length. This is because customers are relatively clustered on these problems.

It is also observed that the total path length becomes better with the increase in the number of customers. The improvement rate shifts from 3.21 % at 50 customers to 9.83 % at 150 customers, indicating that the HFOA and the HAC are more effective when the problems are greater in size. Thus, the HFOA and the HAC are effective means to solve largescale problems, thanks to the local search operations. However, these operations need to consume a certain amount of time.

To verify its efficiency, the HFOA was applied to solve the benchmark problem VRNPC1. The results are displayed in Figure 5. It can be seen that the total path length is progressively shortened from 572.69 unit length, and converges to 551.27 unit length.

Table 3 Simulation results

Problem	Number of customer	Total length of ACS	Total length of HAC	Total length of HFOA	(ACS-HAC)/ACS (%)	(ACS-HFOA)/ACS (%)
Vrpnc1	50	569.564	550.7	551.27	3.31	3.21
Vrpnc2	75	957.525	890.68	886.24	6.98	7.44
Vrpnc3	100	964.132	874.07	870.13	9.34	9.75
Vrpnc4	150	1253.86	1126.12	1130.65	10.19	9.83
Vrpnc5	199	1587.02	1444.29	1440.78	8.99	9.21
Vrpnc6	120	1133.88	1110.45	1109.42	2.07	2.16
Vrpnc7	100	911.861	912.64	910.73	-0.09	0.12
Average		1053.977	986.99	985.60	6.36	6.49

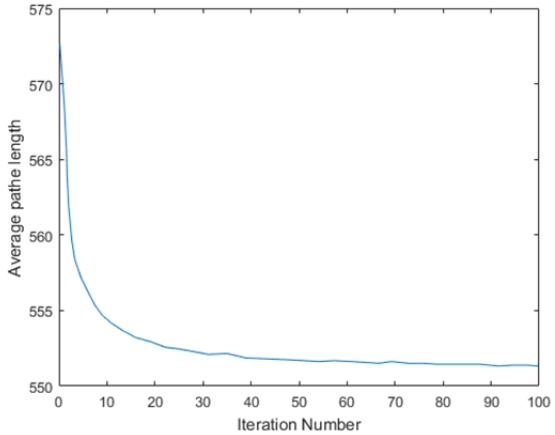


Fig. 5 The results of the HFOA on VRNPC1

4.3 Local search effect

To test the local search effect of HFOA, the three local search operations were implemented to the MCVRP in different combinations. Then, the results were compared with those of the FOA with no local search mechanism. The first combination integrates the FOA with the 2-opt operation, the second combination integrates the FOA with 2-opt and swap, and the third combination integrates the FOA with 2-opt, swap and insertion. Each combination has different number of iterations to achieve an equal computing time. According to the results in Table 4, the FOA with hybrid local search operations is good at shortening the optimal path length of the MCVRP (Fig. 6).

As shown in Table 4, the first combination enhances the solution quality by 11.86 %. However, the effect of the 2-opt operation is not good enough, as it only applies to a single trip. This operation does better in the TSP than the VRP. In the second combination, the solution quality is increased from 11.86 % to 13.04 %, indicating that the swap operation can move the customers on two paths. Finally, the third combination elevates the solution quality from 13.04% to 17.16 %, an evidence of the positive effect of hybrid local search on solution quality.

Table 4 Results of different combinations

Algorithm	Number of iteration	Average path length	%
FOA	10,000	1189.78	-
2-Opt+FOA	10,000	1048.7	11.86
2-Opt+Swap+FOA	300	1034.61	13.04
2-Opt+Swap+Insert+FOA	100	985.60	17.16

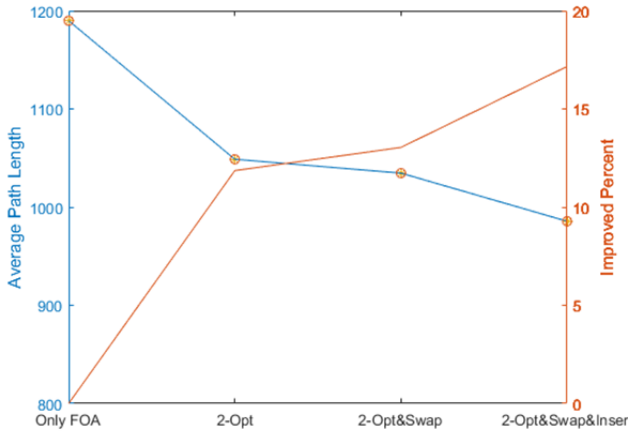


Fig. 6 Results of different combinations

4.4 Multi-compartment vehicle advantages

This section aims to disclose the advantages of the MCV. To do so, the Benchmark problem was solved by two modes. The first mode uses a SCV with the capacity of Q, while the second mode adopts a two-compartment vehicle. The capacity of the two compartments is respectively denoted as Q_1 and Q_2. The two vehicles have the same total capacity, i.e. $Q = Q_1 + Q_2$. Besides, there are two types of products, denoted as 1 and 2 respectively, to be delivered to the customer. Since the two products must be stored separately, a customer can be accessed twice on different paths to transport products 1 and 2. In this case, the VRP is decomposed into two sub-problems. The two sub-problems were solved separately, whose shortest paths were added up to get the result of the first mode. The results of the two modes on the 7 problems are shown in Table 5.

It can be seen from Table 5 that the total path length increases significantly when there is only one compartment. This is because a customer is visited twice to transport the two different products. The repeated visits increase the total travel distance. In the second mode, the two-compartment vehicle can serve more customers on each path than the SCV, thus reducing the total travel length. The advantage of the MCV is demonstrated by the 49.26 % shorter total length of the two-compartment vehicle than the SCV in Table 5.

Table 5 Results of the two modes on the 7 problems

Problem	Number of customer	Total length of vehicle with two compartment (A)	Total length of vehicle with one compartment (B)	(B-A)/A×100
vrpnc1	50	551.27	935.32	69.67
vrpnc2	75	886.24	1319.33	48.87
vrpnc3	100	870.13	1432.56	64.64
vrpnc4	150	1130.65	1719.16	52.05
vrpnc5	199	1440.78	2037.73	41.43
Vrpnc6	120	1109.42	1573.89	41.87
Vrpnc7	100	910.73	1279.78	40.52
Average		985.60	1471.11	49.26

5. Conclusion

The vehicle routing problem is a kind of NP-hard optimization problem with great research significance and wide application value in the intelligent logistics system. Aiming at the problem of vehicle routing optimization with capacity constraints in logistics distribution, a fruit fly optimization algorithm HFOA based on local search is proposed.

Local search operations are fundamental to solving combinatorial optimization problems. In this paper, three local search methods are combined with the FOA into a hybrid optimization algorithm to tackle the MCVRP. The effect of the proposed algorithm was verified with 7 benchmark problems. The numerical experiments show that the HFOA algorithm has improved the performance for all the problems, especially on largescale ones. The author also proved the necessity of the combination of local search and the FOA algorithm, and analysed the advantages of the MCV.

The simulation experiments on seven international benchmark problems show that the effectiveness and stability of the proposed algorithm HILS are compared with other algorithms in the literature. The overall performance of the algorithm HFOA is better. Of course, the HFOA only has a slight lead over the HAC in the total path of the MCVRP. The future research will seek to improve the combination of local search operations and the FOA.

Acknowledgement

This research is supported by the Philosophy and Social Science Foundation of China (#17BGL001), and the Philosophy and Social Science Foundation of Shandong Province (#16CGLJ28).

References

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

[2] Garey, M.R., Johnson, D.S. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman & Co., New York, USA.

[3] Pan, W.-T. (2012). A new fruit fly optimization algorithm: Taking the financial distress model as an example, *Knowledge-Based Systems,* Vol. 26, 69-74, doi: 10.1016/j.knosys.2011.07.001.

[4] Gylienė, V., Eidukynas, V. (2016). The numerical analysis of cutting forces in high feed face milling, assuming the milling tool geometry, *Procedia CIRP,* Vol. 46, 436-439, doi: 10.1016/j.procir.2016.03.132.

[5] Davim, J.P. (2011). *Modern machining technology: A practical guide,* Elsevier, Amsterdam, Netherlands, doi: 10.1533/9780857094940.

[6] Ji, W., Liu, X., Wang, L., Sun, S. (2015). Experimental evaluation of polycrystalline diamond (PCD) tool geometries at high feed rate in milling of titanium alloy TC11, *The International Journal of Advanced Manufacturing Technology,* Vol. 77, No. 9-12, 1549-1555, doi: 10.1007/s00170-014-6517-9.

[7] Mitić, M., Vuković, N., Petrović, M., Miljković, Z. (2015). Chaotic fruit fly optimization algorithm*, Knowledge-Based Systems,* Vol. 89, 446-458, doi: 10.1016/j.knosys.2015.08.010.

[8] Han, J., Liu, C. (2013). Fruit fly optimization algorithm based on bacterial chemotaxis, *Journal of Computer Application*, Vol. 33, No. 4, 964-966, doi: 10.3724/SP.J.1087.2013.00964.

[9] Chajakis, E.D., Guignard, M. (2003). Scheduling deliveries in vehicles with multiple compartments, *Journal of Global Optimization,* Vol. 26, No. 1, 43-78, doi: 10.1023/A:1023067016014.

[10] Avella, P., Boccia, M., Sforza, A. (2004). Solving a fuel delivery problem by heuristic and exact approaches, *European Journal of Operational Research,* Vol. 152, No. 1, 170-179, doi: 10.1016/S0377-2217(02)00676-8.

[11] El Fallahi, A., Prins, C., Calvo, R.W. (2008). A memetic algorithm and a tabu search for the multi-compartment vehicle routing problem, *Computers & Operations Research*, Vol. 35, No. 5, 1725-1741, doi: 10.1016/j.cor.2006.10.006.
[12] Chen, P., Huang, H.-K., Dong, X.-Y. (2009). A multi-operator based iterated local search algorithm for the capacitated vehicle routing problem, *Journal of Beijing Jiaotong University*, Vol. 33, No. 2, 1-5.
[13] Muyldermans, L., Pang, G. (2010). On the benefits of co-collection: Experiments with a multi-compartment vehicle routing algorithm, *European Journal of Operational Research*, Vol. 206, No. 1, 93-103, doi: 10.1016/j.ejor.2010.02.020.
[14] Reed, M., Yiannakou, A., Evering, R. (2014). An ant colony algorithm for the multi-compartment vehicle routing problem, *Applied Soft Computing*, Vol. 15, 169-176, doi: 10.1016/j.asoc.2013.10.017.
[15] Gajpal, Y., Abad, P. (2009). An ant colony system (ACS) for vehicle routing problem with simultaneous delivery and pickup, *Computers & Operations Research*, Vol. 36, No. 12, 3215-3223, doi: 10.1016/j.cor.2009.02.017.
[16] Balseiro, S.R., Loiseau, I., Ramonet, J. (2011). An ant colony algorithm hybridized with insertion heuristics for the time dependent vehicle routing problem with time windows, *Computers & Operations Research*, Vol. 38, No. 6, 954-966, doi: 10.1016/j.cor.2010.10.011.
[17] De la Cruz, J.J., Paternina-Arboleda, C.D., Cantillo, V., Montoya-Torres, J.R. (2013). A two-pheromone trail ant colony system – Tabu search approach for the heterogeneous vehicle routing problem with time windows and multiple products, *Journal of Heuristics*, Vol. 19, No. 2, 233-252, doi: 10.1007/s10732-011-9184-0.
[18] Valíček, J., Harničárová, M., Öchsner, A., Hutyrová, Z., Kušnerová, M., Tozan, H., Michenka, V., Šepelák, V., Mitaľ, D., Zajac, J. (2017). Quantifying the mechanical properties of materials and the process of elastic-plastic deformation under external stress on material, *Materials,* Vol. 8, No. 11, 7401-7422, doi: 10.3390/ma8115385.
[19] Yousefikhoshbakht, M., Didehvar, F., Rahmati, F. (2014). Solving the heterogeneous fixed fleet open vehicle routing problem by a combined metaheuristic algorithm, *International Journal of Production Research,* Vol. 52, No. 9, 2565-2575, doi: 10.1080/00207543.2013.855337.
[20] Liu, Z.-X., Wang, Y.-F., Zhang, Y. (2014). Multiple population fruit fly optimization algorithm for automatic warehouse order picking operation scheduling problem, *Journal of Wuhan University of Technology,* Vol. 36, 71-77.
[21] Karpus', V.E., Ivanov, V.A. (2008). Universal-composite adjustable machine-tool attachments, *Russian Engineering Research,* Vol. 28, No. 11, 1077-1083, doi: 10.3103/S1068798X08110105.
[22] Zheng, X.-L., Wang, L., Wang, S.-Y. (2014). A hybrid discrete fruit fly optimization algorithm for solving permutation flow-shop scheduling problem, *Control Theory & Applications*, Vol. 31, No. 2, 159-164.
[23] Lin, S. (1965). Computer solutions of the traveling salesman problem, *The Bell System Technical Journal,* Vol. 44, No. 10, 2245-2269, doi: 10.1002/j.1538-7305.1965.tb04146.x.
[24] Laporte, G. (1992). The vehicle routing problem: An overview of exact and approximate algorithms, *European Journal of Operational Research*, Vol. 59, No. 3, 345-358, doi: 10.1016/0377-2217(92)90192-C.
[25] Lin, S.-W., Lee, Z.-J., Ying, K.-C., Lee, C.-Y. (2009). Applying hybrid meta-heuristics for capacitated vehicle routing problem, *Expert Systems with Applications*, Vol. 36, No. 2, Part 1, 1505-1512, doi: 10.1016/j.eswa.2007.11.060.
[26] Larbi, A.A., Bounif, A., Bouzit, M. (2018). Comparisons of LPDF and MEPDF for lifted H_2/N_2 jet flame in a vitiated coflow, *International Journal of Heat and Technology*, Vol. 36, No. 1, 133-140, doi: 10.18280/ijht.360118.
[27] Ike, C.C. (2018). Exponential fourier integral transform method for stress analysis of boundary load on soil, *Mathematical Modelling of Engineering Problems,* Vol. 5, No. 1, 33-39, doi: 10.18280/mmep.050105.
[28] Abdulkader, M.M.S., Gajpal, Y., ElMekkawy, T.Y. (2015). Hybridized ant colony algorithm for the multi compartment vehicle routing problem, *Applied Soft Computing,* Vol. 37, 196-203, doi: 10.1016/j.asoc.2015.08.020.

APEM journal
Advances in Production Engineering & Management
Volume 13 | Number 4 | December 2018 | pp 479–491
https://doi.org/10.14743/apem2018.4.305
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Multi-objective transport network design with a reversible simulated annealing algorithm

Feng, X.[a,b,*], **Ruan, Z.**[b], **Zhu, X.**[b], **Zhang, L.**[b]

[a]MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing, P.R. China
[b]School of Traffic and Transportation, Beijing Jiaotong University, Beijing, P.R. China

ABSTRACT

In order to rationally coordinate inconsistent objectives in transport network design, this research newly develops a multi-objective network layout optimisation model solved by an improved Simulated Annealing Algorithm (SAA). Two temperature control variables and one cost difference control variable are defined in the proposed SAA. They work in cooperation to restart the optimum search from the latest temporary optimal solution if the search is made excessively in any searching direction as well as expand the searching area for the globally optimal network layout with the minimum operation cost. The genetic algorithm is embedded into the reversible SAA to iteratively provide a network configuration with the minimum total time expense of all the transports for the minimisation of the network operation cost. It is confirmed that the new optimisation model solved by the reversible SAA integrating the genetic algorithm is able to effectively minimise both the total transport time expense and the network operation cost with searching for the best fits between these two basically inconsistent objectives from different perspectives. The proposed approach can be utilised to optimise configurations of not only urban transit lines for passenger mobility organisation but also logistics transportation routes for manufacturing production management.

ARTICLE INFO

Keywords:
Transport network design;
Multi-objective optimisation modelling;
Reversible simulated annealing algorithm;
Genetic algorithm;
Double temperatures;
Network operation cost difference

**Corresponding author:*
xsfeng@bjtu.edu.cn
(Feng, X.)

Article history:
Received 13 October 2018
Revised 28 November 2018
Accepted 7 December 2018

1. Introduction

Transport network design has been a focused research issue for a long time since the heuristic algorithm proposed by Lampkin and Saalmans [1]. Valuable research achievements have been made from the perspectives of transport efficiency [2, 3], network operation cost [4-6], transport service coverage [4, 7], environment protection [8], etc. Many studies attach importance to the simultaneous minimisations of network operation cost and total transport time expenditure by optimising the layout of transport lines [9-11]. However, because minimising the operation cost of a transport network is, on the whole, inconsistent with saving the time expense of its utilisers the most [10], the optimal solutions are difficult to be satisfyingly found in consideration of different interests of network operator(s) and utilisers [10, 12]. In order to solve this problem, continuous efforts have been made on different aspects.

The most common way to solve the inconsistency of different optimisations is transforming the multi-objective optimisation problem into a mixed-integer programming issue with a linear integration of distinct objectives. For example, the weighted sum of the costs of the operator(s), the users and the unsatisfied travel demands are minimised in the work of Fan *et al*. [13] by op-

timising the bus routes for zonal demand aggregation and variable transit demands. From the perspective of multimodal trips, Cipriani *et al*. [9] try to optimise the configuration of a bus network with elastic travel demands for the minimisation of the linear combination of the operator expense, the user expenditure and the external cost. Similarly taking into account the interests of the operator(s), the users and the external environment, Pternea *et al*. [8] develop an efficient model linearly integrating different objectives to propose a sustainable solution to the transport network design problem. In the research of Chu [11], the weighted sum of the cost for the operations of the bus lines, the generalised expense of all the passengers and the penalty for the unsatisfied travel demands are minimised by simultaneously optimising the bus route network layout and the bus timetables for the network operation.

It is obvious that the weights of various objective functions in the mixed-integer programming studies are difficult to be convincingly determined in a completely rational way. As a result, non-weighted strategies such as alternating, hierarchical and phased optimisation approaches are developed for the multi-objective network configuration optimisation. For instance, in order to minimise both the total time expense of passengers and the amount of necessary vehicles for the operation of a transit network, Arbex and da Cunha [10] develop a new Genetic Algorithm (GA) which is totally different to the conventional one [14, 15]. These two optimisation objectives are cyclically alternated along the generations of the individuals in the newly developed GA. Moreover, in view of the interactions between flows of buses and cars, a bi-level model is proposed by Yu *et al*. [16] to minimise not only the average time cost of travellers taking various travel modes but also the ride comfort difference of passengers on different bus routes in a multi-modal traffic environment for the optimal distribution of bus lanes. From a phased optimising viewpoint, López-Ramos *et al*. [6] develop a lexicographic goal programming model which integrates railway network design phase and train frequency setting phase to minimise the travel time of the passengers as well as the costs for construction and operation of the railway network.

Though effectively applied in some cases, the existing approaches are still not fully able to satisfactorily coordinate the inconsistent objectives in a transport network layout optimisation work, because of various interests of different groups [17]. Inspired by the network topology optimisation studies of Karsten *et al*. [18] and Saad *et al*. [19] in shipping and communication fields, this research focuses on the transport network layout optimisation with a new way of thinking for the optimal coordination of these two inconsistent objectives. Applicable to optimising configurations of both passenger transit lines for mobility organisation and logistics transportation routes for manufacturing production, a multi-objective transport network layout optimisation model is newly developed in this study and solved by an improved Simulated Annealing Algorithm (SAA). The proposed SAA is able to restart the search for the optimum network layout with the minimum operation cost from the latest temporary optimum solution in its iterative executions. Moreover, the GA newly developed by Feng *et al*. [20] is embedded into the reversible SAA to iteratively provide a transport network configuration with the minimal total time cost of all the utilisers to be optimised for the minimisation of the network operation cost.

The remaining parts of this paper are organised as follows. Section 2 first develops the new optimisation model. Thereafter, the SAA proposed in this research is explained in Section 2. Section 3 makes computational experiments and discusses the results to validate the newly developed multi-objective optimisation model solved by the reversible SAA. Finally, Section 4 makes the conclusions of this study and proposes some future research issues.

2. Materials and methods

2.1 Model definition

The stops on the transport network are, in this research, regarded as the points of transport generations and attractions and have sufficient service capacities. In other words, the distance of a transport on the network from the departure place to the starting stop and the distance from the final stop to the destination place are beyond the considerations in this work. Moreover, the

spatial distribution of the transport demands is unchanged. Furthermore, if a stop is passed through by the operating vehicles of a transport line, every vehicle operating on this line provides transport services at this stop. The influences of the dwell time of a vehicle at a stop and the traffic conditions on the network upon the transport time are ignored. That is, the transport time is mainly determined by the transport distance and the average technical Operation Speed (OS) of vehicles in this research. In addition, there are no limitations to the numbers of the operating vehicles of different types for every transport line and all the vehicles operating on a line have the same technical OS in this study.

Furthermore, a transport line is unable to use the vehicles of other lines to meet with its transport demands in this work. A transport line has all of its operating vehicles cyclically start from one of its terminal stops to the other by passing through all the intermediate stops on this line and return by following the same route reversely. One way roads are not considered here. In other words, every two neighbouring stops on a line are connected bi-directionally with two inter-stop transport links consisting of the same successive road links in opposite transport directions. It is also assumed that the preparations of all the vehicles providing transport services on a transport line are completed at one fixed terminal stop of this line.

It is explained by Eq. 1 that one objective of the optimisation model developed in this study is the minimisation of the total operation cost of all the transport lines for a certain time period (e.g., one day) when the transport services are provided. As interpreted by Eq. 2, the operation cost of a transport line is simultaneously decided by multi-factors. The operators need to take into account the cost for operating vehicle maintenance, transport line operation and backup vehicle maintenance. As a result, it is necessary for each line to consider the maintenance cost of each vehicle of every type for operation or backup, the amounts of the vehicles of different types for various purposes and the energy consumption (i.e., usually fuel or electricity cost) of a vehicle of each type for a distance of the transport with some technical OS. The maintenance expense and Energy Cost (EC) intensity of one vehicle of a certain type for some utilisation purpose and the number of the vehicles for backup are all relatively fixed for a transport line. The unit price of the EC for transport operation is also stable in a certain time period. Therefore, determined by the layout of the transport network, both the number of the vehicles operating on each transport line and the transport distance of each operating vehicle mainly decide the operation cost of the network for a transport service time period.

$$Min\ C = Min \sum_{ij} C_{ij} \tag{1}$$

Symbol C represents the operation cost of a transport network, and C_{ij} indicates the operation cost of transport line ij (which provides the circling transport services between the terminal stop i and the terminal stop j) on the network.

$$C_{ij} = \sum_k f_{ij}^k n_{ij}^k + \sum_k c^e g_{ij}^{k,v_{ij}} D_{ij}^k n_{ij}^k + \sum_k f_{ij}^{k,B} n_{ij}^{k,B} \tag{2}$$

The f_{ij}^k is the maintenance cost of one operating vehicle of type k for line ij, n_{ij}^k denotes the number of the operating vehicles of type k for line ij, c^e stands for the unit price of the EC for transport operation, v_{ij} indicates the required technical OS of the vehicles operating on line ij, $g_{ij}^{k,v_{ij}}$ is the EC per unit transport distance (i.e., EC intensity) of one operating vehicle of type k at the technical OS of v_{ij} on line ij, D_{ij}^k denotes the transport distance of an operating vehicle of type k on line ij, $f_{ij}^{k,B}$ represents the maintenance cost of one backup vehicle of type k for line ij, and $n_{ij}^{k,B}$ is the number of the backup vehicle(s) of type k for line ij.

As explained by Eq. 3, the transport distance of an operating vehicle on a transport line is determined by the distance of one entire circle of the transport on the line, the complete circles of the transports provided by the vehicle once it is prepared for its operation and the complete operation times of the vehicle within the operation time period of the transport line. It is interpreted by Eq. 4 that the complete circles of the transports made by a vehicle after its preparation is decided by its Energy Storage Capacity (ESC), EC intensity, Energy Utilisation Rate (EUR) and the distance of one entire circle of the transport on the line. If a vehicle is ready for its operation,

it is able to complete at least one entire circle of the transport on the line where it provides transport services, as explained by Eq. 5.

$$D_{ij}^{k} = d_{ij} c_{ij}^{k} \left[T_{ij}^{O} / \left(\frac{d_{ij} c_{ij}^{k}}{v_{ij}} + t_{ij}^{k,p} \right) \right] \tag{3}$$

In Eq. 3, d_{ij} represents the distance of one entire circle of the transport on line ij, c_{ij}^{k} is the complete circles of the transports provided on line ij by one vehicle of type k once it is prepared for its operation, T_{ij}^{O} stands for the operation time of line ij, and $t_{ij}^{k,p}$ denotes the necessary preparation time of one vehicle of type k for its operation on line ij.

$$c_{ij}^{k} = \left[\left(\left(E^{k} / g_{ij}^{k,v_{ij}} \right) \theta_{ij}^{k,v_{ij}} \right) / d_{ij} \right] \tag{4}$$

In Eq. 4, E^{k} is the ESC one vehicle of type k, and $\theta_{ij}^{k,v_{ij}}$ is the EUR of one vehicle of type k at the technical OS of v_{ij} on line ij.

$$d_{ij} \leq \left(E^{k} / g_{ij}^{k,v_{ij}} \right) \theta_{ij}^{k,v_{ij}} \tag{5}$$

It is indicated by Eq. 6 that the time cost of an entire transport is composed of the time expensed in vehicles operating on different lines and the time consumed in waiting for vehicles. The time used for awaiting a vehicle operating on a transport line is represented here by the maximum headway of all the vehicles operating on this line. Eq. 7 explains the lower limit to the vehicle-kilometers provided by all the vehicles operating on a line. Moreover, the operating vehicles allocated to a transport line should offer the adequate transport capacity for the maximum accumulated transport demand between every two neighbouring stops on this line, as interpreted by Eq. 8. In addition, the upper limit to the time cost of a transport is explained by Eq. 9. The other objective of the optimisation model developed in this research is the minimisation of the total time expense of all the transports, as interpreted by Eq. 10, under the constraints explained by Eq. 5 and Eq. 9.

$$T^{od} = \sum_{ij \in TL^{od}} \left(H_{ij}^{Max} + \left(d_{ij}^{od} / v_{ij} \right) \right) \tag{6}$$

In Eq. 6, T^{od} represents the time expense of a transport from stop o to stop d, TL^{od} denotes all the lines serving the transport from stop o to stop d, H_{ij}^{Max} is the upper limit to the headways of all the vehicles operating on line ij, and d_{ij}^{od} stands for the distance of the transport from stop o to stop d on line ij.

$$d_{ij} \left(T_{ij}^{O} / H_{ij}^{Max} \right) < \sum_{k} D_{ij}^{k} n_{ij}^{k} \tag{7}$$

$$Max \left\{ Max \left\{ \sum_{od} V^{od} \lambda_{pq,Up}^{od}, \sum_{od} V^{od} \lambda_{pq,Down}^{od} \right\}_{pq \in ij} \right\} \leq \sum_{k} U^{k} c_{ij}^{k} \left[T_{ij}^{O} / \left(\frac{d_{ij} c_{ij}^{k}}{v_{ij}} + t_{ij}^{k,p} \right) \right] n_{ij}^{k} \tag{8}$$

Symbol V^{od} denotes the transport demand from stop o to stop d, pq represents the transport section between stop o and stop d which are neighbouring to each other on line ij, $\lambda_{pq,Up}^{od}$ is the 0-1 variable denoting if transport section pq serves the transports on the time-shortest path from stop o to stop d in the upward transport direction (i.e., by taking the value of 1) or not (i.e., by taking the value of 0), $\lambda_{pq,Down}^{od}$ is the 0-1 variable representing if transport section pq serves the time-shortest transports from stop o to stop d in the downward transport direction (i.e., by taking the value of 1) or not (i.e., by taking the value of 0), and U^{k} indicates the carrying capacity of one vehicle of type k.

$$Min\ T^{od} \leq T^{Max} \tag{9}$$

where T^{Max} is the maximum time cost of a transport on the network.

$$Min\ T = Min\ \sum_{od} T^{od} V^{od} \tag{10}$$

where T denotes total transport time cost under constraints interpreted by Eq. 5 and Eq. 9.

2.2 Algorithm design

From the perspective of reducing the operation cost of a transport network the most, transport lines providing circling transport services are built on the shortest paths between every two neighbouring stops and necessary vehicles are allocated to each line, given all the sites of stops. The vehicle allocations need to consider the carrying capacity, OS, ESC, EC intensity, preparation time before operation, etc. of each vehicle, the length of every transport route, the accumulated transport demands between neighbouring stops, etc. However, the transport efficiency in such a case will be very low due to the frequent transfers between every two neighbouring stops. On the contrary, if there is always at least one transport route completely coinciding with the shortest path between any pair of all the stops and adequate vehicles operating on each route do not stop at the intermediate stops they passed through, the total time cost of all the transports will be minimised. Nevertheless, the operation cost of the network at this time will be extremely high because too many vehicles are needed.

It is impossible to construct transport networks in either of these two extreme ways to decrease the network operation cost or increase the transport efficiency. However, the main objective of the network design in one of the extreme cases can be actually transformed into endeavoring to make different lines have no common transport sections but connected with each other through some stops, as shown in Fig. 1. At the same time, from the perspective of minimising the network operation cost, the accumulated transport demands in different transport sections between neighbouring stops of a line need to be similar to each other for avoiding the transport capacity wastes in the transport sections with relatively few accumulations of the transports. In contrast, the objective of building the network in another extreme case can be changed in reality into reducing the time expenses of transports as much as possible in not only riding vehicles but also transfers between transport lines by optimising the layout of the lines for the minimal total time expenditure of all the transports. The network built at this time may be illustrated by Fig. 2. The common transport sections are indispensable in such a case. For instance, the transport section between D and F is commonly used by the line between A and F, the line between B and H, and the line between C and G, as shown in Fig. 2.

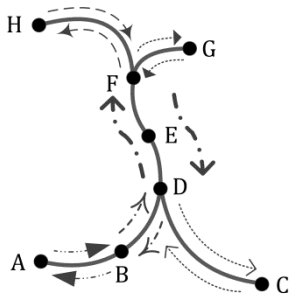


Fig. 1 Example of network layout for relatively low operation cost

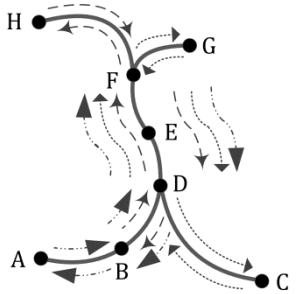


Fig. 2 Example of network layout for less transport time expense

With these two revised objectives, a new SAA explained in Fig. 3 is designed to solve the newly developed multi-objective optimisation model by iteratively searching for the proper network configuration satisfying both of the objectives the most from different viewpoints. Besides the total time cost of all the transports, the expense of the operating vehicles on their transport ser-

vices is also analysed in this work. The initial network layout can be obtained on the basis of the Greedy algorithm [21] or by taking an existing transport network layout directly. Totally different to the ordinary SAA [22], the new SAA proposed in this research defines double temperature control variables (i.e., the iteration temperature control variable and the search temperature control variable). The iteration temperature control variable monitors the reduction of the optimum network operation cost with the iterative computations in the proposed SAA and, at the same time, judges whether the iterative calculations should stop or not. The search temperature control variable takes effect on the probability of adopting an obtained network configuration with a relatively high operation cost as the studied network layout in the iterative execution process of the SAA to explore the optimum solution in various searching directions.

Moreover, the cost difference control variable is also defined to decide the upper limit to the increase of the operation cost of the studied network layout in the executions of the proposed SAA from that of the latest temporally optimal network configuration obtained in a previous iteration of the SAA. If the increase of the network operation cost is no less than the upper limit, the search for the optimum network layout will return to the latest temporally optimal solution, from which, a new search starts. Meanwhile, in view of the relatively strong global optimisation ability [14, 15], the GA proposed by Feng *et al*. [20] is embedded into the reversible SAA to iteratively determine a network configuration with the minimum total time expense of all the transports to be optimised for the minimal operation cost. In this way, the minimisation of the network operation cost and the maximisation of the transport efficiency have the hierarchically alternating coordination to seek their optimal conjunction points.

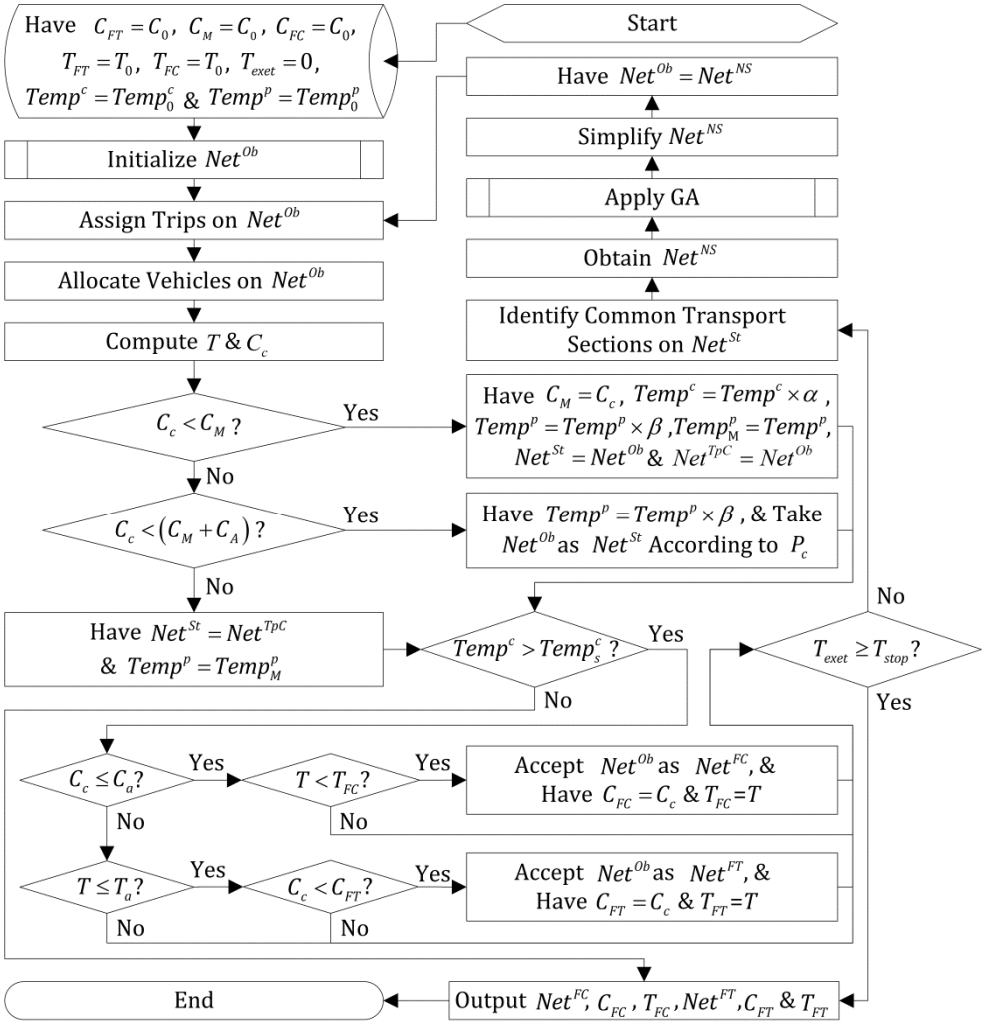


Fig. 3 Optimisation procedure of the proposed SAA

In the optimisation procedure of the proposed SAA, Net^{Ob} and Net^{St} represent the obtained network layout and the studied network layout, respectively. Net^{TpC} is the latest temporally optimal network layout. Net^{FC} denotes the final network layout with the least total transport time in comparison under the constraint of the upper limit to the network operation cost. Net^{FT} is the final network layout with relatively the minimum total operation cost of all the transport lines on the premise of the acceptably low total time expense of all the transports. C_c and C_M are defined as the operation costs of Net^{Ob} and Net^{TpC}, respectively. C_A is the cost difference control variable limiting the maximum value of $(C_c - C_M)$. C_{FC} and C_{FT} are the operation costs of Net^{FC} and Net^{FT}, respectively. C_0 is a very big positive number predetermined as the original value of C_M, C_{FC} and C_{FT}. C_a represents the acceptable maximum network operation cost determined according to transport capacities of all operating vehicles.

Furthermore, T_{FC} and T_{FT} denote the total time costs of all the transports on Net^{FC} and Net^{FT}, respectively. T_0 is another very big positive number predetermined as the original value of both T_{FC} and T_{FT}. T_a indicates the acceptable maximum total transport time. $Temp^c$ denotes the iteration temperature control variable whose original value is $Temp_0^c$. $Temp^p$ is the search temperature control variable whose original value is also $Temp_0^p$. $Temp_s^c$ represents the preset value of $Temp^c$ for stopping the iterative executions of the algorithm. $Temp_M^p$ denotes the specific value of $Temp^p$ at which Net^{TpC} is obtained. The values of $\alpha \in (0,1)$ and $\beta \in (0,1)$ are predetermined to control the decreasing speeds of $Temp^c$ and $Temp^p$, respectively. P_c represents the probability of adopting Net^{Ob} as Net^{St} when the operation cost of Net^{Ob} is more than that of Net^{TpC}. T_{exet} is the time spent on the iterative executions of the proposed SAA, and T_{stop} indicates the upper limit to the accumulated time for executing the reversible SAA designed in this research. The detailed steps of the optimisation process of the proposed SAA are as follows:

Step 1: Have $T_{exet} = 0$, $C_M = C_0$, $C_{FC} = C_0$, $C_{FT} = C_0$, $T_{FC} = T_0$, $T_{FT} = T_0$, $Temp^c = Temp_0^c$ and $Temp^p = Temp_0^p$, and initialise Net^{Ob}.

Step 2: Assign all the transport demands to each of their time-shortest transport routes on Net^{Ob}, according to Eq. 6. If two or more time-shortest transport routes from one stop to another are found, the corresponding transport demand is equally split to each of the shortest routes.

Step 3: Allocate necessary vehicles to each transport line on Net^{Ob} to satisfy not only Eq. 7 but also Eq. 8. Compute the total time cost of all the transport demands on Net^{Ob} (i.e., T) and C_c.

Step 4: If $C_c < C_M$, have $C_M = C_c$, $Temp^c = Temp^c \times \alpha$, $Temp^p = Temp^p \times \beta$ and $Temp_M^p = Temp^p$, and adopt Net^{Ob} as not only Net^{St} but also Net^{TpC}.
If $C_M \le C_c < (C_M + C_A)$, keep $Temp^c$, Net^{TpC}, $Temp_M^p$ and C_M unchanged, have $Temp^p = Temp^p \times \beta$, and decide whether or not to take Net^{Ob} as Net^{St}, according to P_c calculated by Eq. 11.

$$P_c = exp((C_M - C_c)/Temp^p) \times 100\ \% \tag{11}$$

If $(C_M + C_A) \le C_c$, have $Net^{St} = Net^{TpC}$ and $Temp^p = Temp_M^p$.

Step 5: If $Temp^c > Temp_s^c$, $C_c \le C_a$ and $T < T_{FC}$, accept Net^{Ob} as Net^{FC}, and have $C_{FC} = C_c$ and $T_{FC} = T$.
If $Temp^c > Temp_s^c$, $T \le T_a$ and $C_c < C_{FT}$, accept Net^{Ob} as Net^{FT}, and have $C_{FT} = C_c$ and $T_{FT} = T$.

Step 6: If $Temp^c \le Temp_s^c$ or $T_{exet} \ge T_{stop}$, stop the executions of this algorithm, and output Net^{FC}, C_{FC}, T_{FC}, Net^{FT}, C_{FT} and T_{FT}.

Step 7: Identify each common transport section on Net^{St} for every aggregation of the transport lines with two common sets of the successive inter-stop road links in reverse transport directions. The identifications follow the decreasing order of the quantities of the transport lines aggregated by each of the common transport sections. Moreover, every two of the identified common transport sections have not any same part.

Step 8: Build transport lines based on each of the identified common transport sections on Net^{St} to obtain a new network layout (i.e., Net^{NS}).

Step 9: Apply the GA to update the layout of the lines on Net^{NS} for the minimisation of the total transport time expenditure on satisfying all the transport demands, as explained by Eq. 10, under the constraints interpreted by both Eq. 5 and Eq. 9.

Step 10: Delete each line coinciding completely with (a part of) another one to simplify the updated Net^{NS}. Take Net^{NS} as Net^{Ob}, and return to Step 2.

In the applied GA, the scale of the population is fixed and the chromosome amount of an individual is unchanged. Each individual (i.e., the transport network layout) is initialised with all the chromosomes (i.e., the transport lines) which are sequenced in the same order in different individuals. Every chromosome sequences its genes (i.e., the stops), according to the order of each stop on a transport line. Each gene can be found in at least one chromosome of every individual. Mutation is implemented to a certain ratio (e.g., Q %) of all the individuals. After mutating every randomly chosen chromosome of each individual belonging to the Q % of the population by probabilistically changing the configuration of the transport network on different aspects, the transport demands between each pair of the stops are assigned on every network layout to the time-shortest transport paths, according to Eq. 6. If there are two or more time-shortest paths from one stop to another, the corresponding transport demand is split equally to each of them. (100.00 % – Q %) of all the individuals with relatively small T are reserved in selection. The rest individuals (i.e., Q % of the population) are randomly paired to carry out the crossovers swapping probabilistically chromosomes at the same positions. Thereafter, the mutation is implemented again. Such iterative executions of the GA do not stop until a network configuration which is able to satisfy all the transport demands with the minimal T is obtained for the network operation cost minimisation in the proposed SAA.

3. Results and discussion

Fig. 2 provides the road network, illustrates the stops (i.e., A, B, C, D, E, F, G and H) and presents the initial layout of the transport network. The transport distances between neighbouring stops are explained in Table 1. The initial network layout consists of the lines with transports between each pair of A, B, D, E and F, every two of B, D, E, F and H, and all the pairs of C, D, E, F and G, respectively. Two computational experiments are made for the symmetrical and non-symmetrical distributions of the same total transport demand, which are correspondingly explained in Table 2 and Table 3. In order to simply the experimental computations, all the transport lines are supposed to have the same operation time period and use the same electric operating vehicles whose carrying capacity, ESC, EC intensity, EUR, preparation time before operation, etc. are determined in advance. Moreover, vehicles operating on different lines are supposed to have not only the same headway but also the same technical OS. Furthermore, it is also assumed that there is no upper limit to the time expense of a transport in the computational experiments. Values of some key parameters and variables are shown in Table 4 for the computational experiments. If $(C_c - C_M)$ in last iteration of executing the proposed SAA is negative, C_A in current execution of the SAA has a very big positive value. If $(C_c - C_M)$ in not only last and but also current execution of the proposed SAA is positive or zero, C_A in current iteration of executing the reversible SAA takes $\lambda(C_c - C_M)$ of last iterative execution of the SAA. The value of λ is $1/3$ here.

Table 1 Transport distances between neighbouring stops

Stop	A	B	C	D	E	F	G	H
A	-	10.00	-	-	-	-	-	-
B	10.00	-	-	8.00	-	-	-	-
C	-	-	-	15.00	-	-	-	-
D	-	8.00	15.00	-	6.00	-	-	-
E	-	-	-	6.00	-	9.00	-	-
F	-	-	-	-	9.00	-	11.00	13.00
G	-	-	-	-	-	11.00	-	-
H	-	-	-	-	-	13.00	-	-

Table 2 Symmetrical transport demand matrix

Stop	A	B	C	D	E	F	G	H	Sum by Origin
A	0	400	700	200	300	450	220	390	2660
B	400	0	360	410	260	500	480	320	2730
C	700	360	0	240	630	280	420	290	2920
D	200	410	240	0	390	430	230	310	2210
E	300	260	630	390	0	330	270	440	2620
F	450	500	280	430	330	0	420	370	2780
G	220	480	420	230	270	420	0	550	2590
H	390	320	290	310	440	370	550	0	2670
Sum by Destination	2660	2730	2920	2210	2620	2780	2590	2670	21180

Table 3 Non-symmetrical transport demand matrix

Stop	A	B	C	D	E	F	G	H	Sum by Origin
A	0	473	705	222	340	522	240	404	2906
B	317	0	390	482	278	552	547	336	2902
C	695	330	0	250	692	286	426	325	3004
D	178	338	230	0	447	454	244	349	2240
E	260	242	568	333	0	381	278	514	2576
F	378	448	274	416	279	0	496	398	2689
G	200	413	414	216	262	344	0	654	2503
H	376	304	255	271	366	342	446	0	2360
Sum by Destination	2404	2548	2836	2190	2664	2881	2677	2980	21180

Table 4 Values of parameters and variables

Parameters and Variables	Values
Carrying Capacity	40 passengers per vehicle
Technical OS	30.00 kilometers per hour
EC Intensity	1.00% of ESC per kilometer
EUR	90.00% of ESC
Preparation Time	2.50 hours
Upper Limit to Headways	0.50 hours
Transport Line Operation Time	10.00 hours
α	0.97
β	0.94
C_A	$+\infty$ or 1/3 of last $(C_c - C_M)$

As to the symmetrical transport demands explained in Table 2, it is shown in Fig. 4 that the decreasing trends of both the iteration temperature (i.e., $Temp^c$) and the optimum network operation cost (i.e., C_M) with the iterative executions of the proposed SAA are completely the same with each other. When the iteration temperature keeps invariant, the decrease of the search temperature (i.e., $Temp^p$) and the corresponding increase of the studied network operation cost (i.e., C_c) are reversible in different optimisation searching directions within a certain limit determined by the cost difference control variable. Moreover, as shown in Fig. 5 and Fig. 6, with the overall decrease of C_c, the total time expense of all the transports (i.e., T) generally keeps increasing, and a decrease of T is usually accompanied by a corresponding increase of C_c, because of their inconsistency. On the premise that the network operation cost (i.e., C_{FC}) is no more than the acceptable maximum value (i.e., C_a), the network layout (i.e., Net^{FC}) with the least total time expense (i.e., T_{FC}) is indicated to be one solution for the symmetrical transport demands, as illustrated in Fig. 6. In contrast, under the constraint of the upper limit (i.e., T_a) to the total time expense (i.e., T_{FT}), another solution is the network configuration (i.e., Net^{FT}) with the least operation cost (i.e., C_{FT}). It is also clarified in Fig. 6 that, only at the beginning of executing the proposed SAA, reducing C_c and saving T are consistent with each other. Represented by Net^C in Fig. 6, the network layout with both C_c and T minimised at such time can be specified by the executive iterations of the proposed SAA as the solution for the consistent minimisations of C_c and T, if C_c or/and T of Net^C is/are correspondingly no more than C_a or/and T_a.

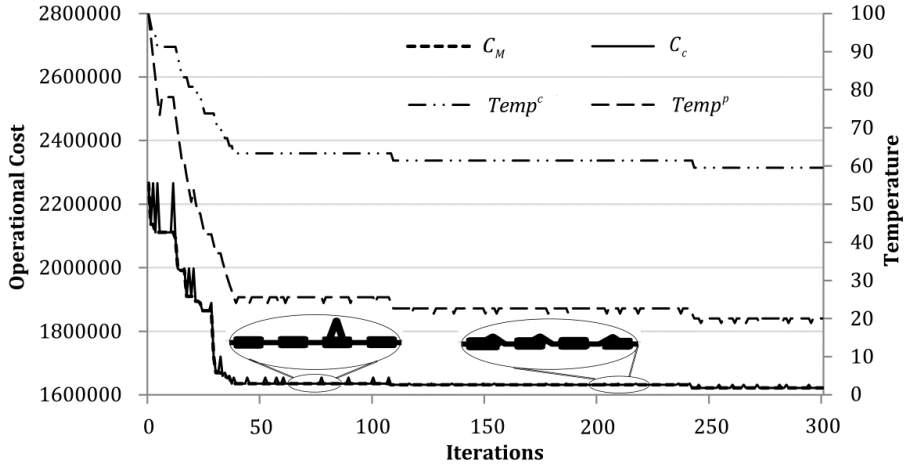


Fig. 4 Changes of C_M, C_c, $Temp^c$ and $Temp^p$ for the symmetrical transport demands

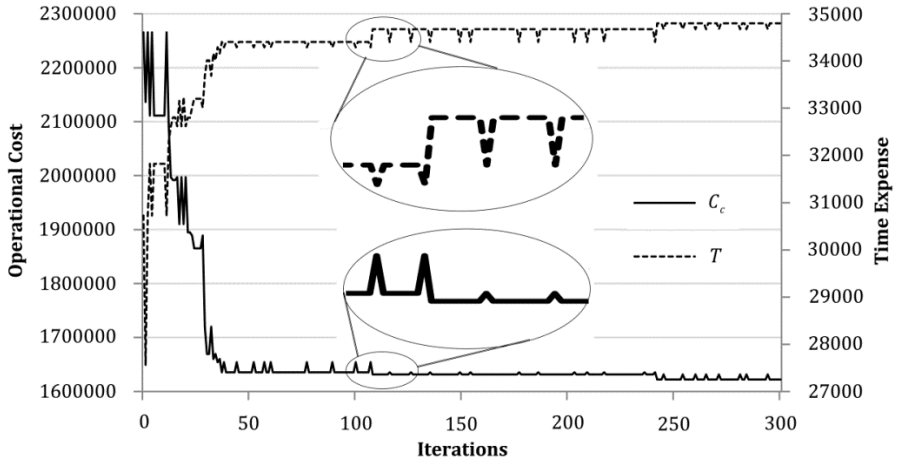


Fig. 5 Changes of C_c and T for the symmetrical transport demands

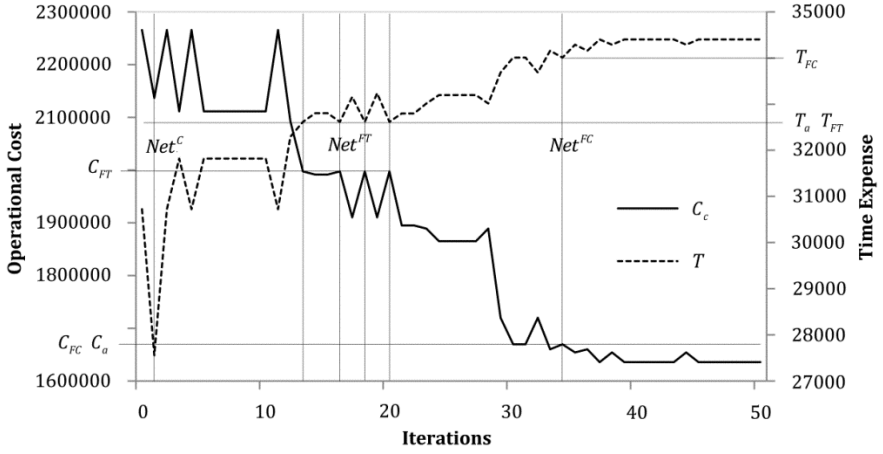


Fig. 6 Examples of the solutions for the symmetrical transport demands

In terms of the non-symmetrical transport demand distribution interpreted in Table 3, the changes of C_M, C_c, $Temp^c$, $Temp^p$ and T with the iterative executions of the reversible SAA are completely similar to their variations for the symmetrical demands, as explained in Fig. 7, Fig. 8 and Fig. 9. Due to the relative complexity of minimising both C_c and T for the non-symmetrical demands, C_M and $Temp^p$ decrease slower in comparison to their decreases with the iterations of executing the proposed SAA for the symmetrical demands. It is apparent that both minimising the operation cost of all the transport lines and maximising the efficiency of all the transports can still be effectively achieved in coordination from different viewpoints.

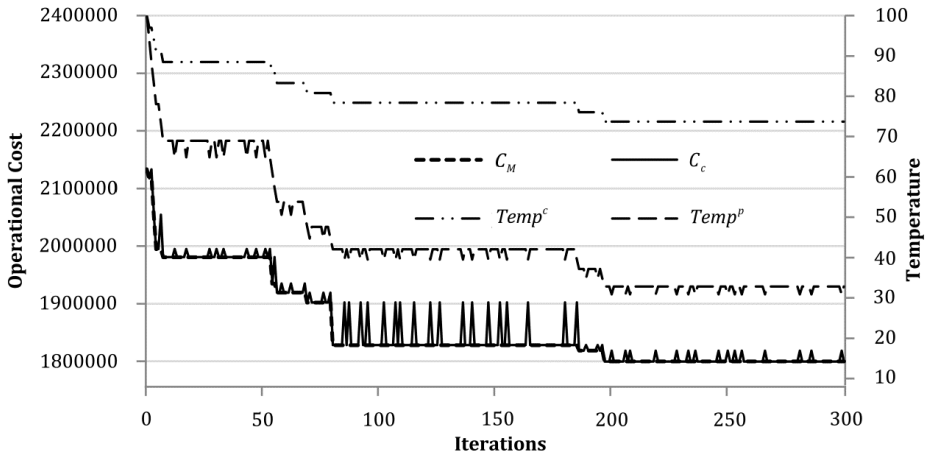


Fig. 7 Changes of C_M, C_c, $Temp^c$ and $Temp^p$ for the non-symmetrical transport demands

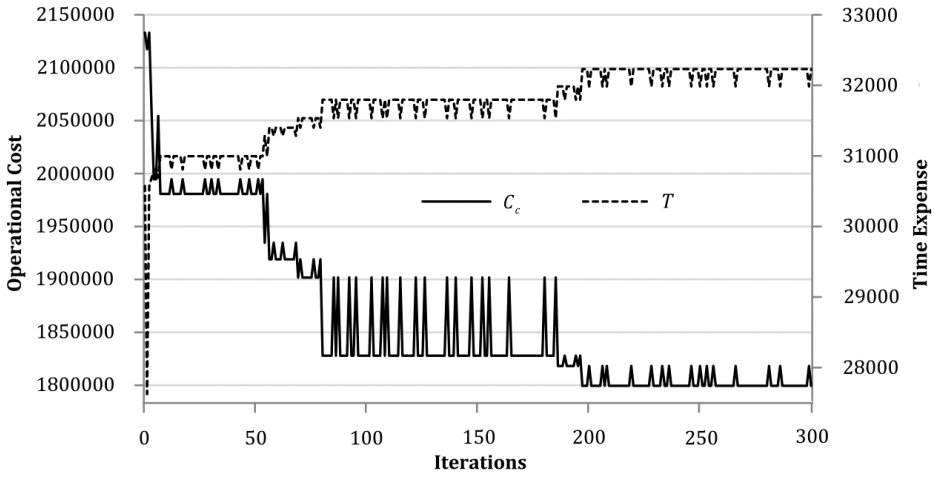


Fig. 8 Changes of C_c and T for the non-symmetrical transport demands

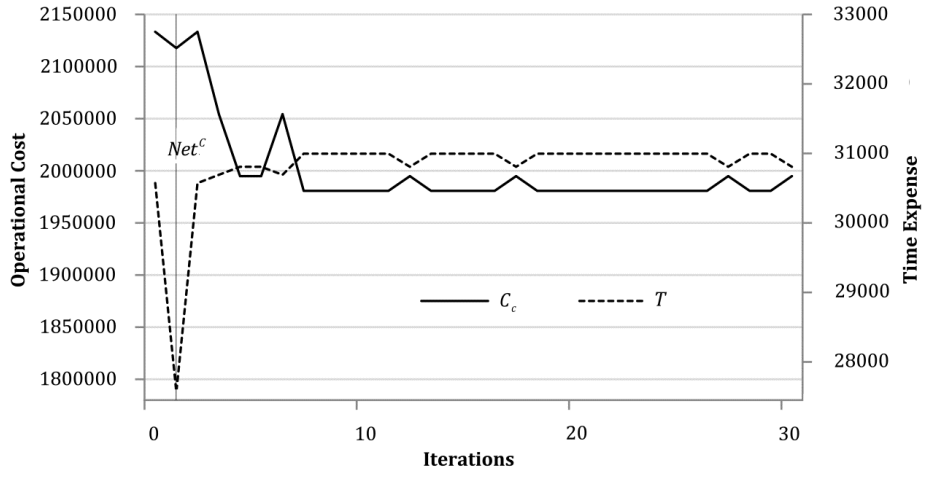


Fig. 9 Possible solution for the non-symmetrical transport demands

4. Conclusion

This research develops a new multi-objective transport network layout optimisation model solved by a newly proposed reversible SAA with the GA embedded. The area of searching for the optimum solution in iteratively executing the proposed SAA is rationally expanded by double temperature control variables. Moreover, the cost difference control variable makes the network

layout optimisation able to restart from the latest temporally optimal solution by stopping an excessive search in any searching direction to avoid getting trapped in a local optimum. These three control variables work together to improve the capacity of the new SAA for the search of the globally optimal solution. The GA integrated into the reversible SAA iteratively decides a basic transport network configuration with the minimum total transport time to be optimised for the minimisation of the network operation cost. It is confirmed that the proposed model solved by the reversible SAA with the GA embedded is able to effectively optimise the layout of a passenger transit network or configurations of logistics transportation routes for both improving the transport efficiency and reducing the operation cost with the best fits between them.

In future research, the newly developed optimisation model solved by the proposed SAA ought to have more values of its different parameters and variables tested for different networks with various topologies to further validate the achievements of this study. Moreover, the dynamic designs of the transport services should be made in consideration of real-time traffic condition, changeable transport demand distribution, sharing vehicles between different transport lines, vehicle maintenance, etc. in future work. Furthermore, instead of controlling the minimisation of the network operation cost, the best fits between the maximisation of the transport efficiency and the minimisation of the network operation cost might also be achieved coordinately by controlling the search for the optimum total time expense of all the transports. This is worthy of further analyses in detail. In addition, as proposed by Kılıç and Gök [17] and Cheng *et al*. [23], respectively, both rationally initializing the configuration of a transport network with a relatively high efficiency and hierarchically optimising the network layout from environmental protection perspective are another two important tasks in the future.

Acknowledgements

This study is supported by National Natural Science Foundation of China [grant number 71571011] and the Fundamental Research Funds for the Central Universities [grant number 2018JBM022].

References

[1] Lampkin, W., Saalmans, P.D. (1967). The design of routes, service frequencies, and schedules for a municipal bus undertaking: A case study, *Journal of the Operational Research Society*, Vol. 18, No. 4, 375-397, doi: 10.1057/jors.1967.70.

[2] Nikolić, M., Teodorović, D. (2013). Transit network design by bee colony optimisation, *Expert Systems with Applications*, Vol. 40, No. 15, 5945-5955, doi: 10.1016/j.eswa.2013.05.002.

[3] Nayeem, M.A., Rahman, Md.K., Rahman, M.S. (2014). Transit network design by genetic algorithm with elitism, *Transportation Research Part C: Emerging Technologies*, Vol. 46, 30-45, doi: 10.1016/j.trc.2014.05.002.

[4] Nikolić, M., Teodorović, D. (2014). A simultaneous transit network design and frequency setting: Computing with bees, *Expert Systems with Applications*, Vol. 41, No. 16, 7200-7209, doi: 10.1016/j.eswa.2014.05.034.

[5] Larrain, H., Muñoz, J.C., Giesen, R. (2015). Generation and design heuristics for zonal express services, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 79, 201-212, doi: 10.1016/j.tre.2015.04.008.

[6] López-Ramos, F., Codina, E., Marín, Á., Guarnaschelli, A. (2017). Integrated approach to network design and frequency setting problem in railway rapid transit systems, *Computers & Operations Research*, Vol. 80, 128-146, doi: 10.1016/j.cor.2016.12.006.

[7] Gutiérrez-Jarpa, G., Laporte, G., Marianov, V., Moccia, L. (2017). Multi-objective rapid transit network design with modal competition: The case of Concepción, Chile, *Computers & Operations Research*, Vol. 78, 27-43, doi: 10.1016/j.cor.2016.08.009.

[8] Pternea, M., Kepaptsoglou, K., Karlaftis, M.G. (2015). Sustainable urban transit network design, *Transportation Research Part A: Policy and Practice*, Vol. 77, 276-291, doi: 10.1016/j.tra.2015.04.024.

[9] Cipriani, E., Gori, S., Petrelli, M. (2012). A bus network design procedure with elastic demand for large urban areas, *Public Transport*, Vol. 4, No. 1, 57-76, doi: 10.1007/s12469-012-0051-7.

[10] Arbex, R.O., da Cunha, C.B. (2015). Efficient transit network design and frequencies setting multi-objective optimisation by alternating objective genetic algorithm, *Transportation Research Part B: Methodological*, Vol. 81, Part 2, 355-376, doi: 10.1016/j.trb.2015.06.014.

[11] Chu, J.C. (2018). Mixed-integer programming model and branch-and-price-and-cut algorithm for urban bus network design and timetabling, *Transportation Research Part B: Methodological*, Vol. 108, 188-216, doi: 10.1016/j.trb.2017.12.013.

[12] Guihaire, V., Hao, J.-K. (2008). Transit network design and scheduling: A global review, *Transportation Research Part A: Policy and Practice*, Vol. 42, No. 10, 1251-1273, doi: 10.1016/j.tra.2008.03.011.

[13] Fan, W.D., Machemehl, R.B. (2008). Some computational insights on the optimal bus transit route network design problem, *Journal of the Transportation Research Forum*, Vol. 47, No. 3, 61-75, doi: 10.5399/osu/jtrf.47.3.2116.
[14] Goldberg, D.E. (1989). *Genetic algorithms in search, optimisation, and machine learning*, Addison-Wesley Longman Publishing, USA.
[15] Sivanandam, S.N., Deepa, S.N. (2008). *Introduction to genetic algorithms*, Springer-Verlag, Berlin, Heidelberg, Germany.
[16] Yu, B., Kong, L., Sun, Y., Yao, B., Gao, Z. (2015). A bi-level programming for bus lane network design, *Transportation Research Part C: Emerging Technologies*, Vol. 55, 310-327, doi: 10.1016/j.trc.2015.02.014.
[17] Kılıç, F., Gök, M. (2014). A demand based route generation algorithm for public transit network design, *Computers & Operations Research*, Vol. 51, 21-29, doi: 10.1016/j.cor.2014.05.001.
[18] Karsten, C.V., Brouer, B.D., Desaulniers, G., Pisinger, D. (2017). Time constrained liner shipping network design, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 105, 152-162, doi: 10.1016/j.tre.2016.03.010.
[19] Saad, A., Khan, S.A., Mahmood, A. (2018). A multi-objective evolutionary artificial bee colony algorithm for optimising network topology design, *Swarm and Evolutionary Computation*, Vol. 38, 187-201, doi: 10.1016/j.swevo.2017.07.010.
[20] Feng, X., Zhu, X., Qian, X., Jie, Y., Ma, F., Niu, X. (2018). A new transit network design study in consideration of transfer time composition, *Transportation Research Part D: Transport and Environment*, Article in press, doi: 10.1016/j.trd.2018.03.019.
[21] Kolo, B.A. (2009). *Introduction to matroids*, Weatherford Press, Weatherford, USA.
[22] Chong, E.K.P., Zak, S.H. (2013). *An introduction to optimisation,* Fourth edition, John Wiley & Sons, Inc., Hoboken, USA.
[23] Cheng, H., Madanat, S., Horvath, A. (2016). Planning hierarchical urban transit systems for reductions in greenhouse gas emissions, *Transportation Research Part D: Transport and Environment*, Vol. 49, 44-58, doi: 10.1016/j.trd.2016.08.033.

APEM journal
Advances in Production Engineering & Management
Volume 13 | Number 4 | December 2018 | pp 492–503
https://doi.org/10.14743/apem2018.4.306
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Compatibility of ionic liquids with hydraulic system components

Kambič, M.[a,*], **Kalb, R.**[b], **Tič, V.**[c], **Lovrec, D.**[c]

[a]Olma d.o.o., Ljubljana, Slovenia
[b]Proionic GmbH, Grambach, Austria
[c]University of Maribor, Faculty of Mechanical Engineering, Maribor, Slovenia

ABSTRACT

The aim of this work was to identify, which of the known ionic liquids used within the technical area, primarily as a lubricant, would also be appropriate for use as a hydraulic fluid. In this context, their suitability has been proved based on experimental research with respect to the appropriate physical and chemical properties as required for mineral based hydraulic fluid. Primary aim of the research was to determine the ability of ILs to protect against corrosion, which is one of the important factors in choosing an ionic liquid. The results show that, despite excellent lubricating properties, certain ILs fail on this corrosion test. Except the corrosion protection performance of the basic hydraulic components parts, e.g. hydraulic pumps and valves, in the foreground was their compatibility with other materials used within other parts of hydraulic system, e.g. coating of the hydraulic tank and the filter material. For this purpose standard tests methods for mineral based hydraulic oils have been used, supplemented with non-standard tests, carried out at the same conditions as they occur during the operation of the hydraulic system.

ARTICLE INFO

Keywords:
Ionic liquids;
Hydraulic fluid;
Corrosion protection;
Material compatibility

**Corresponding author:*
milan.kambic@olma.si
(Kambič, M.)

Article history:
Received 13 June 2018
Revised 25 November 2018
Accepted 29 November 2018

1. Introduction

Mineral based hydraulic oil is nowadays still the most widely used type of hydraulic fluid, with all its advantages and disadvantages in regard to its specific field of application. There is no ideal solution, so we are tending to attain them as closely as possible, always searching for new options resp. alternatives. One of the alternative hydraulic fluids is water (pure water), which is in fact environmentally acceptable, low-cost, non-flammable fluid with low isothermal compressibility. But due to the other major disadvantages, such as poor lubrication properties, low viscosity, tendency to corrosion, the consequent problems need to be solved by using specific (more expensive) materials and manufacturing processes or including special coatings. And at the end, it can only be used under limited operating conditions [1-3]. All these drawbacks limit the overall use of water as a hydraulic fluid, preventing a direct replacement of mineral hydraulic oil with water.

The newer, very promising option to a future alternative to today's commonly used hydraulic fluids represents ionic liquids (ILs). Based on their unconventional properties, ionic liquids allow fundamentally new approaches to technical challenges. They have the potential to open doors to radical innovations [4].

Ionic liquids were first reported as very promising high-performance lubricants in 2001 and have attracted considerable attention within the field of tribology since then because of their

remarkable lubrication and anti-wear capabilities as compared with lubrication oils in general use. Ionic liquids have many of good features as described in different literature. Therefore they should be the ideal candidates for new lubricants, suitable for use under harsh conditions, where conventional oils and greases or solid lubricants fail [5-8].

A large number of studies have already been carried out in this area so far but only a few with ionic liquids suitable for use within hydraulic systems. Despite their excellent individual properties, it is very difficult to find an ionic liquid that would combine the majority of good characteristics, essential for use as a hydraulic fluid [9-11].

Besides the low isothermal compressibility, as an excellent property of the most ionic liquids [12], its good lubricating property is among the next more important conditions for the use of ionic liquids as a hydraulic fluid. Their lubricating properties need to be better or at least the same as of conventional hydraulic mineral based oil. This problematic has so far been quite well studied by a number of authors (e.g. [13-19]) and is not in the forefront of this paper. However, good lubricating properties serve for the pre-selection of appropriate ionic liquids for further testing [19].

In addition to good lubricating properties, the ionic liquids as an appropriate medium for the use in the high-pressure hydraulic system must also have other good properties. For wide commercial use they need to meet numerous special requirements, e.g. harmlessness to human health, good thermal and chemical stability, low corrosion impact on the metallic parts of hydraulic components, and be compatible with various materials used within hydraulic components and system, such as seals, coating protection, and the filter material. The latter properties we will focus on in the next sections.

2. Materials and methods

Good corrosion protection is in addition to the good lubrication properties the next very important feature of a hydraulic fluid. Due to the fact that the ionic liquids are actually salts, can be expected that the corrosion protection capacity would be one of the parameters, the most difficult to approach the properties of the conventional hydraulic liquids, particularly, the mineral-based oils. That was confirmed in laboratory test, particularly, in corrosion test in a humid chamber, where most of tested ionic liquids proved to be considerably worse than the mineral hydraulic oil.

Compared to the other corrosion test, e.g. by using the rotating cage (e.g. [20]), the test in humid chamber is very rigorous and very useful test for a rapid assessment the corrosion protection of a different hydraulic fluids. Because of very tightened-up conditions, the time until occurrence of corrosion it is very short. Besides humid chamber test, the corrosion protection capacity was determined by the standard method of determination of corrosiveness to copper and practical method of determination of corrosion in the open air. Corrosion tests were so performed in three different ways:

- Testing of corrosion-preventing in a condensation water alternating atmosphere according DIN 51386-1 [21]. Corrosion test in humid chamber (DIN EN ISO 6270-2) [22] was conducted at constant conditions. The chamber was closed throughout the test, with temperature of 40 °C and relative humidity of 100 %).
- Standard test method for corrosiveness to copper from petroleum products by copper strip test according ASTM D 130-04 (copper is a commonly used material for sliding surfaces inside hydraulic pumps and motors).
- Corrosion in open air (practically method, in accordance with the real operating atmosphere conditions).

2.1 Corrosion test in humid chamber

The corrosion test in a humid chamber is very suitable for comparing different protection materials with known corrosion resistance and those in the course of testing. The corrosion protection capacity in humid chamber is determined in detail by the standard DIN EN ISO 6270-2. Test-

ing is carried out at constant atmosphere in duration until the occurrence the first signs of corrosion and sometimes also after that in order to gain additional information.

The humid chamber is a closed container the bottom of which is covered with heated distilled water. In that way, the relative humidity in the chamber is always 100 %. The tests are often performed with interruptions. Table 1 shows the individual testing circumstances as defined by that standard.

For the test in the humid chamber the apparatus type HK 400 and accessories have been used, and as a testing material steel metal plates (according to standard) made from steel type ST 1405 (according to DIN 1623, part 1), with thickness of 0.7 mm to 0.8 mm, roughness of R_a = 0.6 µm to R_a = 1.9, and the standard plate size 150 × 210 mm. The test plates were prior the testing sanded with the 120-grit sandpaper, and cleaned with acetone and ethanol.

The humid chamber must be placed in a space without aggressive atmosphere, with room temperature and relative air humidity of not more than 75 %. It must be so positioned that it is protected against air draught and direct sun rays. The bottom of chamber has to be filled with distilled water between 10 and 12 mm level. The water level must be checked each time prior to heating. The test plates have to be placed within the test space at the same level as shown in the Fig. 1.

The polished, degreased and dry sheet steel plate should be thinly coated with protective agent or lubricant (in our case with mineral based hydraulic oil or ionic fluid) having to completely cover the cleaned surface, and then introduced into the humid chamber. The required temperature in the chamber in amount of 40 °C ± 3 °C must be reached within 90 minutes, condensed water having to form on the samples.

During the test in constant atmosphere (K), the constant temperature (40 °C ± 3 °C) must be maintained throughout the testing. Testing has to be effected in so many cycles that the first visible signs of corrosion – brown stains (1 cycle = 24 hours) appear. When testing in an alternating condensation atmosphere (A1), heating must be interrupted after 8 hours, reckoning from the heating activation, the humid chamber door opened and left open for 16 hours. In this way, the first cycle is at an end. Then, the distilled water level has to be checked, water added if necessary, and the device closed. Other details regarding the test procedure and testing samples are given in the standards DIN EN ISO 6270-2.

Table 1 Atmospheric types in corrosion tests in a humid chamber [21]

Test atmosphere	Code	Cycle duration (h)	Test period (min, h)	Air temperature (°C)	Relative humidity (%)
Constant	K	-	-	40 ± 3	100
Alternating condensation atmosphere	A1	24	8 h (including warm-up)	40 ± 3	100
			16 h (including cooling down, chamber open)	18 to 28	< 100
	A2	24	8 h (including warm-up)	40 ± 3	100
			16 h (including cooling down)	18 to 28	100

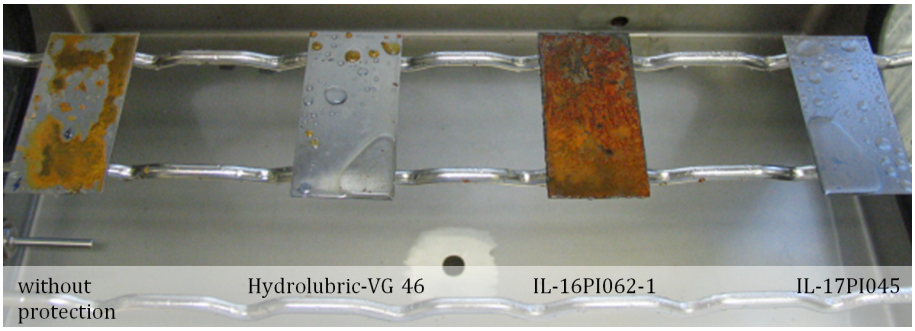


Fig. 1 Test plates coated with different lubricants in humid chamber after 24 h

2.2 Corrosiveness to copper

Because of the presence of non-ferrous metals based on copper in hydraulic components the test of corrosiveness to copper also was carried out. The corrosiveness to copper is carried out in accordance to standard ASTM D 130 [23].

The polished copper plate has to be dipped into the sample (50 ml) at a prescribed temperature and time. The test tube has to be placed into the bath at the specified temperature 50 °C or 100 °C and specified time of heating (usually for 3 h). After three hours, heating must be stopped, the plate taken out of the test tube and corrosiveness assessed with the standard copper plates (etalons), as shown in Fig. 7. Corrosiveness is expressed by values 1 to 4 obtained by comparing corroded copper plate with freshly polished standard plate by assessing the appearance of the test plate (ASTM Copper Strip Corrosion Standards).

2.3 Corrosion in the open air at the ambient condition

In this case a non-standardized, practically test was performed. Identical steel plates as for the humid chamber corrosion test were used. Also the plate preparation was the same. Each plate was coated with the ionic liquid sample, whereby possible changes of the surface appearance were observed at the ambient conditions (room temperature approximately about 20 °C and the humidity about 60 %).

2.4 Compatibility with hydraulic components materials

Presented standard tests of the corrosion protection in a humid chamber, corrosiveness to copper and practically test of corrosion in the open air, were performed with the standardised test materials, as described in the previous chapter. In addition to these test it is reasonable to check compatibility the ILs with the materials that are used for the manufacturing of hydraulic components – pumps and especially valves. Valve housing, valve control piston, springs, washers, etc. are the most exposed vital part of a hydraulic component and thus the entire hydraulic system.

For practically testing the compatibility with these materials the more interesting ionic liquids IL-17PI045 (as very promising IL) and EMIM-$EtSO_4$ (as often mentioned and researched IL) were used (these ILs have been chosen on the basis of extensive IL pre-selection research [24, 25], as explained in chapter 2; see also chapter 3). Our target was to test the effect of both ionic liquids on the component parts the real hydraulic systems.

For the purpose of this test the original components of the renowned hydraulic equipment producer have been procured. They are shown in Fig. 2. The following components were used:

- directional control valve piston made of steel 10715, case hardened 0.4 mm and hardened to 58-62 HRC (Fig. 2, item 1),
- spring made of steel spring wire according to DIN 17223-1 (Fig. 2, item 2),
- washer made of steel DC01 (Fig. 2, item 3),
- bolt made of steel 10715 without subsequent thermal treatment (Fig. 2, item 4),
- valve housing made of grey cast iron GG30, DIN 1691, stress-relief annealed (Fig. 2, item 5),
- valve housing made of steel 10718, drilled from blank (Fig. 2, item 6),
- valve housing made of steel Hyt 60R (Fig. 2, item 7),
- piston RTB 15 made of steel 16MnCr5 (Fig. 2, item 8),
- piston housing RTB 15 made of steel 42CrMo4 without subsequent thermal treatment (Fig. 2, item 9),
- piston VE 60 made of steel 16MnCr5, case hardened (Fig. 2, item 10).

Prior to the test start, the component parts were not specially prepared, but were used such as supplied by the producer. After completion of manufacture, they were protected by him with low-viscosity corrosion protection oil. They were used in such a condition as they are usually, used in practice. Exceptions are all three valve housings from which small samples were cut off by saw.

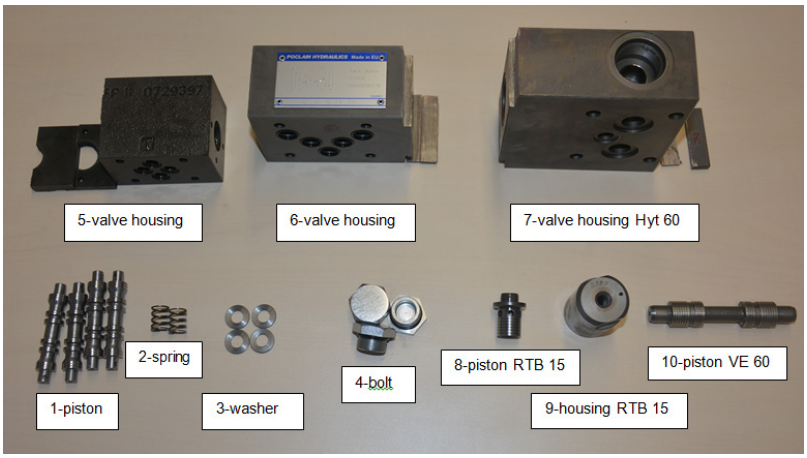


Fig. 2 Samples of hydraulic valve parts

The compatibility test was so performed in this way that different components were first fully wetted and then permanently dipped in the samples of ionic liquids EMIM-$EtSO_4$ and IL-17PI045. Some of the component parts, which are also installed in the same valve, were dipped in the same glass cup. In that way, items 1, 2 and 3 shown in Fig. 2, were united and so were the items 5, 6 and 7 and items 9 and 10. In the liquid IL-EMIM-$EtSO_4$ items 1, 2, 3, 4, 5, 6 and 7, but not the items 8, 9 and 10 were dipped. In the liquid IL-17PI045 all items from 1 to 10 were dipped. Glass cups with test components prior to the test start are shown in Fig. 3.

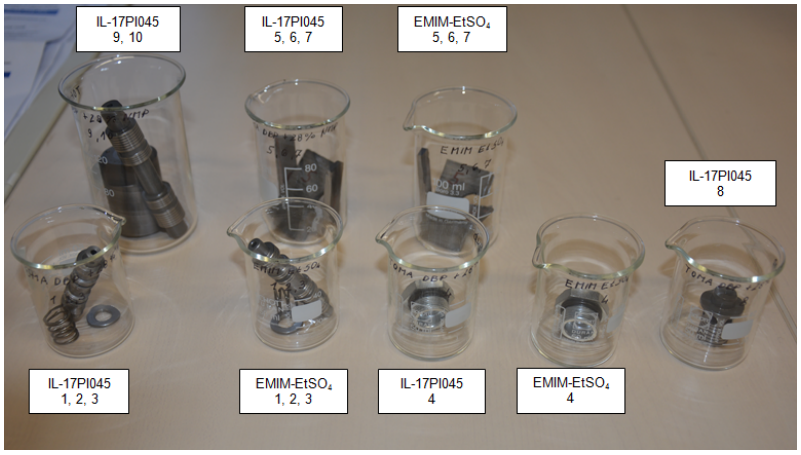


Fig. 3 Hydraulic valve samples prior to soaking

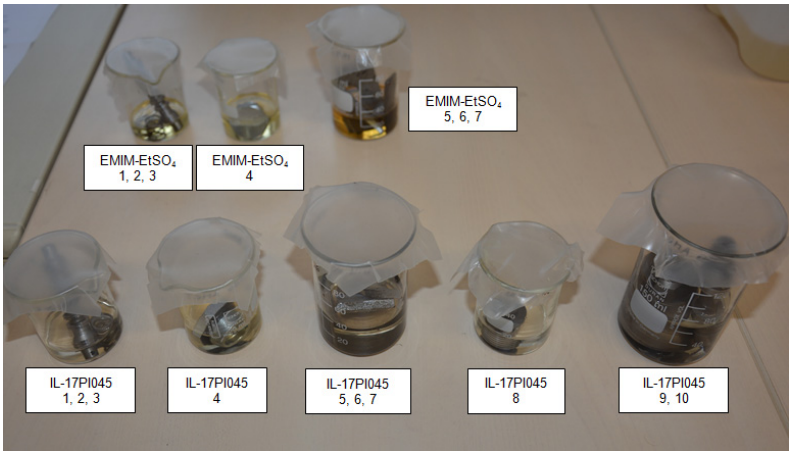


Fig. 4 Hydraulic valve samples during test of compatibility with ionic liquids

After the insertion the components parts into each cup, they were poured with 10 to 30 ml of ionic liquid. The components parts were so firstly wetted with the ionic liquid, and then placed into upright position, so that the bottom part was completely immersed, while the remaining part above the surface level, during testing, was exposed to the influence of room temperature and humidity (with the exception of washer, which was completely immersed). This corresponds to the actual situation in the hydraulic component when the hydraulic system is not under pressure (standby mode) or when the components are temporarily stored. Fig. 4 shows the condition during execution of compatibility test.

2.5 Compatibility with paint coats

Most hydraulic tanks are painted inside and outside, with different types of paint coats. Therefore, it is very important that the ionic liquid used for purposes as a hydraulic liquid should be compatible with the used paint coat. As known, the compatibility the paint coats with mineral based hydraulic oils is not a problem.

Concerning protective coatings for hydraulic reservoirs, in particular the compatibility of the protective paint in regard to the type of hydraulic fluid, there is not a specific standard which would be related to this issue in detail. There are more or less recommendations or recommended practices (RP), mostly linked to related areas, providing guidance on achieving effective corrosion control in storage tanks. They contain information pertinent to the selection of lining materials, surface preparation, lining application, cure, and inspection of tank bottom linings for existing and new storage tanks, e.g. API RP 652 Standard - Linings of Aboveground Petroleum Storage Tank Bottoms. Thus, manufacturers of lubricants are using simple, practical experiments, e.g. testing by continuous contact through immersing of painted metal samples in the liquid under test, at constant room temperature: 20 to 25 °C.

The test of paint coat compatibility was performed with both selected ionic liquids. The metal plate samples were painted with two paints typically used for the tank interior and exterior. The interior is painted with epoxy type priming coat, while the exterior is additionally coated with epoxy type thick-layer finishing coat. Some metal plate samples were painted only with the interior paint and some were additionally coated still with the paint for the tank exterior.

The test was performed so that about 40 ml of each of the ionic liquid were poured into glass cups; then single metal plate samples coated with the paint coat for the tank interior and a metal plate sample coated with the paint coat for the tank exterior were introduced into them. At the beginning of testing the entire metal plates were wetted with the test liquid; during testing the bottom part of the metal plate was permanently dipped into the liquid, while the top part was above the surface level as shown in Fig. 5.

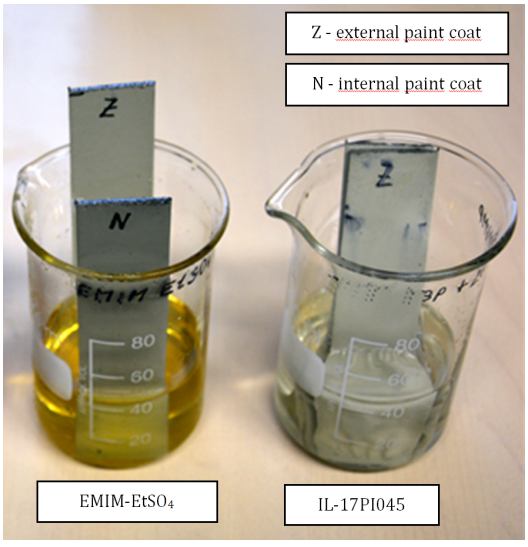


Fig. 5 Test of compatibility with paint coats

2.6 Compatibility with filter material

As the liquid in the hydraulic system acts as a lubricant reducing wear of hydraulic elements, reduction of solid particles in the circulation is important. This is particularly applicable, when clearances in the hydraulic system are small, as, then, a high degree of liquid cleanliness is required for trouble-free operation and reaching the expected useful life time of components. Contaminants are removed by filters. Therefore, the next relevant issue is the hydraulic fluid compatibility with the filter cartridge material.

The ability of the hydraulic liquid to flow through fine filter elements without their clogging is called the filterability or filtration capacity. The filtration capacity of mineral oils is assessed by the laboratory test method according to standard ISO 13357 [26]. With respect to mentioned the compatibility with the filter material, can be carried out in the very close connection with the filterability test.

The first part of standard ISO 13357 covers the testing in the presence of water in oil and the second part the testing without it. For other liquids, in principle, the test is not applicable because the testing filters membranes may be are not compatible with them. According to that standard the membrane filter from mixed cellulose esters, of 47 mm diameter and pore size 0.8 µm is to be placed in a Petri dish which, in turn, is to be put into the oven with 70 °C ± 2 °C temperature for 10 minutes.

We wanted to determine in the described manner also the filterability of both samples of ionic liquids EMIM-$EtSO_4$ and IL-17PI045 as well the mineral oil Hydrolubric VG 46 and at the same time the compatibility with the cellulose as a filter material.

As has already been indicated in the Standard ISO 13357, that in addition to mineral oils the standardised test method can be used also for other liquids, but the latter are maybe not compatible with membrane filters, in our case that proved to be true. As both ionic liquids deformed the cellulose filter membrane and the filterability determination by that method was not possible. The filter element covered with IL-17PI045 started to decompose immediately after the contact with liquid and was dissolved. The sample of EMIM-$EtSO_4$ ionic liquid deformed the filter element into a glassy circle.

Taking into account the negative effect of ionic liquids IL-17PI045 and EMIM-$EtSO_4$ on the cellulose filter paper when trying to determine the filterability according to standard ISO 13357, an additional analysis was performed by us, too. The 0.45 µm filter paper of type ME 25 (Schleicher & Schuel) and 0.8 µm filter paper of type ME 27 (Whatman) were used for the test of compatibility of the mentioned liquids with the cellulose filter paper.

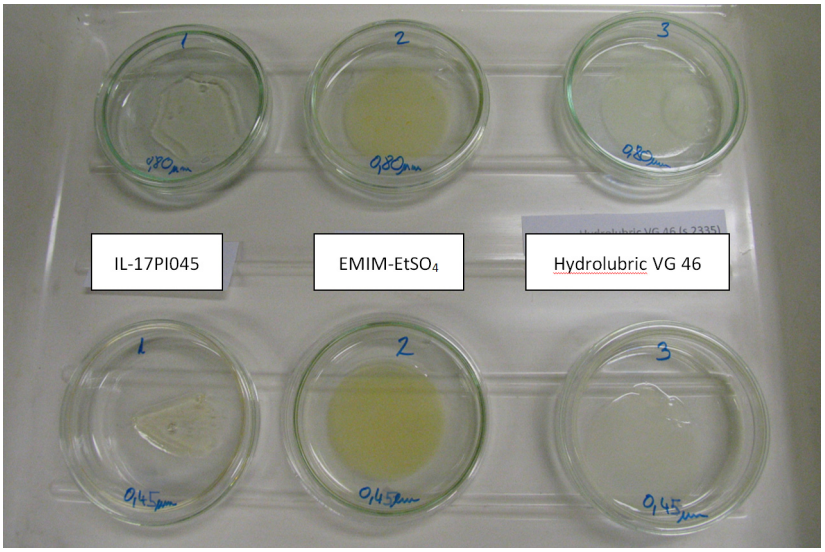


Fig. 6 Compatibility of ionic liquids of mineral oil with cellulose filter paper

3. Results and discussion

3.1 Corrosion test in humid chamber

According to the described method, a larger number of ILs has been tested. Types of testing ILs were carefully selected on the basis of practical experiences from related fields [4] and according to their tribological properties identified by the previous tribological tests – welding point and wear diameter test [19, 24]. The pre-selection procedure can be carried out also in the reverse order - firstly the corrosiveness test and then the lubrication properties test.

Table 2 shows time to first occurrence of corrosion for only some of ILs, in comparison to mineral hydraulic oil Hydrolubric VG 46. It can be seen that corrosion protection in most cases does not amount to 1 cycle (24 hours). In that case, the time in minutes or hours up to the occurrence of the first signs of corrosion is stated in parentheses. In most cases, the corrosion protection in comparison with the mineral hydraulic oil is worse – shaded in gray.

In other cases, the corrosion protection capacity is comparable to that of the mineral hydraulic oil or even much better, as it can be seen on the last three samples in the Table 2 respectively from Fig. 1. According to the viscosity suitable for use in most hydraulic systems, particularly, the last two samples (IL-10 and IL-11) are very interesting. Beside the appropriate viscosity they have higher welding load point, smaller wear diameter (which corresponds to the lower sliding friction between component parts and lower wear of the component parts), and much higher viscosity index (the smaller change of viscosity with temperature). Otherwise the best results of the corrosion test in humid chamber out of all ionic liquids were measured on the sample IL-17PI064 (quaternary ammonium dialkylphosphate + 7 % H_2O) which was the only one to offer protection against corrosion for 1 cycle (24 hours) and/or even slightly longer. However, the viscosity of that sample is too high for most hydraulic systems.

Table 2 Comparison of some physical-chemical properties

Property/Method	Welding point IP 239-85 [kg]	Wear diameter IP 239-85 [mm]	Viscosity 40 °C ASTM D 445 [mm^2/s]	Viscosity index ASTM D 2270 [/]	Corrosion in Humid chamber DIN EN ISO 6270-2
Sample					
Hydrolubric VG 46,	140	0.58	47.07	119	0 (3 h)
IL EMIM-$EtSO_4$	180	1.00	39.44	168	0 (15 min)
IL-EMIM-TFSI	1120	0.65	71.89	132	0 (1.5 h)
IL-10PI462 (EMIM-TFSI)	-	-	-	-	0 (30 min)
IL-16PI028-5 (quaternary ammonium dialkylphosphate)	-	-	-	-	0 (30 min)
IL-10PI028-3 (quaternary ammonium perfluorocarboxylate)	-	-	-	-	0 (45 min)
IL-16PI062-2 (quaternary ammonium dialkylphosphate)	135	0.82	59.14	-	0 (4 h)
IL-16PI062-1 (quaternary ammonium perfluorocarboxylate)	-	-	61.46	-	0 (5 min)
IL-18PI094 (quaternary ammonium perfluorocarboxylate + 30 % EG)	125	1.04	49.28	109	0 (15 min)
IL-17PI064 (quaternary ammonium dialkylphosphate +7 % H_2O)	190	0.49	102.90	105	1
IL-18PI163 (quaternary ammonium dialkylphosphate + 40 % NMP)	160	0.38	47.36	155	0 (3.5 h)
IL-17PI045	145	0.35	46.59	155	0-1 (> 7.5 h)

3.2 Corrosiveness to copper

The results are given as the corrosiveness to Cu with the designation of the corrosion degree determined from the table or by comparison with the etalon, as shown in Fig. 7, by stating the temperature and duration of the test.

The test was carried out (according to ASTM D 130) with two ionic liquid samples, IL-17PI045, as the most promising for the use as hydraulic fluid and EMIM-$EtSO_4$ as often referred and used within different technical application.

After completion of testing, the appearance of the copper strip was compared with the etalon as shown in Fig. 7. For both tested ILs there were no visible changes, meaning that the result of that test is 1a: both ionic liquids are compatible with materials containing copper. To the same conclusion we have come with other tested ILs – corrosiveness to copper is much lower than that of steel and does not represent any problems.

Fig. 7 Determination of corrosiveness to copper for IL-11 17PI045

3.3 Corrosion in the open air at the ambient condition

In case of mineral hydraulic oil the corrosion in the open air did not appear even after a long time period (more than 60 days), while in case of some ionic liquids it appeared already after 20 to 30 minutes. In case of ionic liquid EMIM-$EtSO_4$ the corrosion occurred too, already after 2 to 3 days. Fig. 8 shows the condition after 4 days since the test start. It can be observed that the condition is worse on the plate bottom part, where the ionic liquid layer was slightly thicker due to uneven base.

In case, when the steel plate, after coating with ionic liquid IL-17PI045, was left in the open air at room temperature, corrosion did not appear even after 3 months since the test start. Therefore the testing in the open air was interrupted after this period.

Fig. 8 Test of corrosion in the open air for EMIM-$EtSO_4$ after 4 days

3.4 Compatibility with hydraulic components materials

The test showed compatibility of hydraulic valve components with both interesting ionic liquids. During several-month test period the changes of component part surface were not observed. This was surprising, particularly, when using the ionic liquid EMIM-$EtSO_4$ which did not witness good corrosion protection either in the corrosion test in humid chamber or in the test of corrosion in the open air. A possible reason could be that in this case the component parts were previously protected by a purpose anticorrosion agent. On item 5 sawn from the valve housing and not protected on the sawn surface, the first signs of corrosion did appear after about 90 days in case of EMIM-$EtSO_4$, as shown in Fig. 9.

According to the results of the compatibility test corrosion problems are not to be expected in the practical use of both ionic liquids in the hydraulic system, unless water is present in the system.

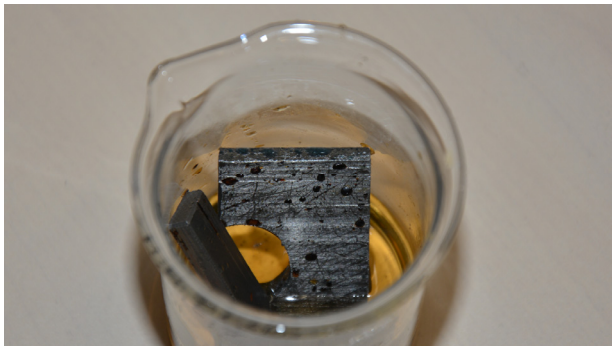


Fig. 9 Test of soaking in EMIM-$EtSO_4$ after 90 days; item 5

3.5 Compatibility with paint coats

Already after a few days of testing, it was proved that the ionic liquid IL-17PI045, which in all previous tests it proved to be an excellent substitute to mineral oil was not compatible either with the paint coat for the tank interior or with the coat for the tank exterior. The paint coat starts to wrinkle and peeled off on the part of the metal plate permanently dipped in the test liquid. The condition after 4 days of the test of this IL is shown in Fig. 10.

The ionic liquid EMIM-$EtSO_4$ is compatible with the usual paint coats of the tank metal plates and does not present any problem in this regard. After more than 7 months of testing, no changes of the condition of painted metal plate occurred either on the permanently dipped part of the metal plate or on the wetted part above the surface level.

The results of this test confirm the generally fact that (nowadays) there is no an ideal hydraulic fluid, which also applies for the ionic liquid. In the case of certain superior properties for a particular IL may be other properties conversely poor. Thus, in the case of IL-17PI045 as an excellent candidate for the future hydraulic fluid, the latter problem can be solved by another kind of protection or other materials for the hydraulic tank.

Fig. 10 Ionic liquid IL-11 (17PI045) and painted metal plate

3.6 Compatibility with filter material

The filter paper was covered with a thin layer of mentioned liquids and their effect was observed (Fig. 6). After 5 minutes, the filter paper covered with ionic liquid IL-17PI045 shrivelled and after 15 minutes it became gelatinous. The ionic liquid IL-1 (EMIM-$EtSO_4$) and the mineral hydraulic oil did not have such an effect on the filter paper. The filter paper covered with EMIM-$EtSO_4$ became only a little brittle.

According to the results of testing of the filterability in conformity with and compatibility with cellulose filter paper discussed in this point, it can be concluded that the ionic liquids IL-17PI045 and EMIM-$EtSO_4$ are not compatible with the cellulose filter paper. Therefore, the use of cellulose filter elements within hydraulic systems with those two liquids is not recommendable. But this incompatibility with the cellulose does not represent a major practical problem for the filtering of ionic liquids within the hydraulic system, as they are for the filter cartridges different materials available, e.g. micro-fiberglass or polyester.

The problem represents the filterability test according the ISO 13357 standard, that dictates the cellulose. The performing the filterability test need the reconsider the use of other filter membrane materials, for example glass fibres. This would, however, still have to be established.

4. Conclusion

The presented results of the research work allow the suggestions for the use of ionic liquids within hydraulic systems and some limitations. Most ionic liquids tested were corrosive in the presence of moisture. Consequently, even greater attention must be paid to moisture prevention in the hydraulic system. Alternatively, stainless hydraulic components can be used, which, however, results in undesirable structural and price changes.

Some ionic liquids, such as, e.g. IL-17PI045, shows the best proposition for use as a hydraulic fluid, since it has much better properties than mineral oil. But on the other hand is not compatible with conventional paint coats of hydraulic tanks. Therefore, the latter must not be painted, when that liquid is used. Another limitation in the use of this liquid is incompatibility with the cellulose filter elements. As a result, absorption filter elements, usually based on cellulose, could not be utilized, when this liquid is used in the hydraulic system.

At present, one of the greatest limitations for wider technical use of ionic liquids, also in hydraulic systems, is at the moment (due to the production of small quantities) considerably higher price than that of conventional hydraulic mineral oil. The target areas of use of ionic liquids in hydraulic equipment are currently the small-volume hydraulic systems, such as wind turbines, mobile hydraulic equipment etc.

The performed researches are only a modest beginning of researches of this extensive area of new lubricants with promising properties. In particular, in the continuation of the research work it would be reasonable to test the selected ionic liquids still in the real hydraulic system. All the obtained results of the presented researches and the mentioned limitations can be of assistance in selecting proper system components and represent guidelines for further work in this sphere.

Acknowledgement

This research was supported by company Proionic GmbH from Grambach/Austria which provided all the samples of ionic liquids and was very cooperative partner in IL-selection process, for sharing their wisdom with us. We are also grateful to the company OLMA d.o.o. from Ljubljana (Slovenia), which allowed us to use their equipment and facilities, as well as the personnel to carry out very extensive experimental work. We are thankful to all colleagues in both companies who provided their expertise and skills that greatly assisted this research.

References

[1] Majdič, F., Kalin, M. (2014). Test rig and comparison of pressure changes at transient phenomena in water- and oil-based power-control hydraulics, *Journal of Vibroengineering*, Vol. 16, 401-411.

[2] Majdič, F., Kalin, M. (2014). Characteristics of the stationary behaviour of water- and oil-based power-control hydraulics, *Mechanika*, Vol. 20, No. 3, 274-281, doi: 10.5755/j01.mech.20.3.5301.

[3] Majdič, F., Velkavrh, I., Kalin, M. (2013). Improving the performance of a proportional 4/3 water-hydraulic valve by using a diamond-like-carbon coating, *Wear*, Vol. 297, No. 1-2, 1016-1024, doi: 10.1016/j.wear.2012.11.060.
[4] Kalb, R. (2015). Ionic liquids – A new generation of high-tech engineering liquids, In: *Proceedings of international conference Fluid Power 2015*, University of Maribor, Slovenia, 49-77.
[5] Zhou, F., Liang, Y., Liu, W. (2009). Ionic liquid lubricants: Designed chemistry for engineering applications, *Chemical Society Reviews*, Vol. 9, No. 38, 2590-2599, doi: 10.1039/b817899m.
[6] Kondo, Y., Yagi, S., Koyama, T., Tsuboi, R., Sasaki, S. (2012). Lubricity and corrosiveness of ionic liquids for steel-on-steel sliding contacts, *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, Vol. 226, No. 11, 991-1006, doi:10.1177/1350650112456127.
[7] Jiang, D., Hu, L., Feng, D. (2011). Crown-type ionic liquids as lubricants for steel-on-steel system, *Tribology Letters*, Vol. 41, No. 2, 417-424, doi: 10.1007/s11249-010-9726-x .
[8] Fox, M.F., Priest, M. (2008). Tribological properties of ionic liquids as lubricants and additives, Part 1: Synergistic tribofilm formation between ionic liquids and tricresyl phosphate, *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, Vol. 222, No. 3, 291-303, doi: 10.1243/13506501JET387.
[9] Amiril, S.A.S., Rahim, E.A., Syahrullail, S. (2017). A review on ionic liquids as sustainable lubricants in manufacturing and engineering: Recent research, performance, and applications, *Journal of Cleaner Production*, Vol. 168, 1571-1589, doi: 10.1016/j.jclepro.2017.03.197.
[10] Regueira, T., Lugo, L., Fernández, J. (2014). Ionic liquids as hydraulic fluids: Comparision of several properties with those of conventional oils, *Lubrication Science*, Vol. 26, No. 7-8, 488-499, doi: 10.1002/ls.1235.
[11] Nyberg, E., Respatiningsih, C.Y., Minami, I. (2017). Molecular design of advanced lubricant base fluids: Hydrocarbon-mimicking ionic liquids, *RSC Advances*, Vol. 7, No. 11, 6364-6373, doi: 10.1039/c6ra27065d.
[12] Kambic, M., Kalb, R., Tasner, T., Lovrec, D. (2014). High bulk modulus of ionic liquid and effects on performance of hydraulic system, *The Scientific World Journal*, Vol. 2014, Article ID 504762, doi: 10.1155/2014/504762.
[13] Pensado, A.S., Comuñas, M.J.P., Fernández, J. (2008). The pressure-viscosity coefficient of several ionic liquids, *Tribology Letters*, Vol. 31, No. 2, 107-118, doi: 10.1007/s11249-008-9343-0.
[14] Minami, I. (2009). Ionic liquids in tribology, *Molecules*, Vol. 14, No. 6, 2286-2305, doi: 10.3390/molecules 14062286.
[15] Van Rensselar, J. (2010). Cover story: Unleashing the potential of ionic liquids, *Tribology and Lubrication Technology*, Vol. 66, No. 4, 24-31.
[16] Ye, C., Liu, W., Chen, Y., Yu, L. (2001). Room-temperature ionic liquids: A novel versatile lubricant, *Chemical Communications*, No. 21, 2244-2245, doi: 10.1039/b106935g.
[17] Jiménez, A.-E., Bermúdez, M.-D. (2007). Ionic liquids as lubricants for steel-aluminium contacts at low and elevated temperatures, *Tribology Letters*, Vol. 26, No. 1, 53-60, doi: 10.1007/s11249-006-9182-9.
[18] Pejaković, V., Kronberger, M., Mahrova, M., Vilas, M., Tojo, E., Kalin, M. (2012). Pyrrollidinum sulfate and ammonium sulfate ionic liquids as lubricant additives for steel/steel contact lubrication, *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, Vol. 226, No. 11, 923-932, doi: 10.1177/1350650112448978.
[19] Kambić, M., Kalb, R., Lovrec, D. (2015). Lubrication properties of ionic liquids suitable for use within hydraulic systems, In: *Proceedings of International Conference Fluid Power 2015*, University of Maribor, Slovenia, 79-93.
[20] Uerdingen, M., Treber, C., Balser, M., Schmitt, G., Werner, C. (2005). Corrosion behaviour of ionic liquids, *Green Chemistry*, Vol. 7, No. 5, 321-325, doi: 10.1039/b419320m.
[21] Standard DIN 51386-1 (1986). Testing of corrosion preventive oils in a condensation water alternating atmosphere.
[22] Standard ISO 6270-2:2005 (2005). Paints and varnishes – Determination of resistance to humidity – Part 2: Procedure for exposing test specimens in condensation-water atmospheres.
[23] Standard ASTM D130-12 (2012). Standard test method for corrosiveness to copper from petroleum products by copper strip test.
[24] Kambič, M., Kalb, R. (2013). Corrosion and lubrication properties of some lonic liquids, *Ventil*, Vol. 19, No. 5, 364-368.
[25] Lovrec, D. (2017). Ionic liquid as a novel, high performance hydraulic fluid – Selection process, In*: ISCAME 2017: Proceedings, 5th International Scientific Conference on Advances in Mechanical Engineering*, University of Debrecen, Hungary, 300-305.
[26] Standard ISO 13357-2:2005 (2005). Petroleum products – Determination of the filterability of lubricating oils - Part 2: Procedure for dry oils.

Calendar of events

- International Conference on Industrial Engineering and Engineering Management, December 16-19, 2018, Bangkok, Thailand.
- 11th International Conference on Computer Modeling and Simulation, January 16-19, 2019, Melbourne, Australia.
- International Conference on Green Manufacturing and Production Engineering, January 17-18, 2019, Rome, Italy.
- 16th Annual Congress on Materials Research and Technology, February 18-20, 2019, Amsterdam, Netherlands.
- 3rd International Conference on 3D Printing Technology and Innovations, March 25-26, 2019, Rome, Italy.
- International Conference on Intelligent Manufacturing and Intelligent Materials, May 9-11, 2019, Sanya, China.
- International Conference on Agile and Flexible Manufacturing Systems, May 21-22, 2019, Berlin, Germany.
- International Conference on Advanced Manufacturing Technologies and Intelligent Machining, May 29-30, 2019, Osaka, Japan.
- 17th Annual Industrial Simulation Conference, June 5-7, 2019, Lisbon, Portugal.
- International Conference on Design and Implementation of Intelligent Manufacturing Systems, June 10-11, 2019, Tokyo, Japan.
- 6th International Conference and Exhibition on Automobile & Mechanical Engineering, July 8-9, 2019, Zurich, Switzerland.
- 23rd International Conference on Advanced Materials & Nanotechnology, August 19-20, 2019, Tokyo, Japan.
- AI Manufacturing 2019: Machine Learning and Artificial Intelligence, The Fourth Industrial Revolution, August 28-29, 2019, Westin O'Hare, Rosemont, Illinois, USA.
- 9th IFAC Conference on Manufacturing Modeling, Management, and Control, August 28-30, 2019, Berlin, Germany.

Notes for contributors

General

Articles submitted to the *APEM journal* should be original and unpublished contributions and should not be under consideration for any other publication at the same time. Manuscript should be written in English. Responsibility for the contents of the paper rests upon the authors and not upon the editors or the publisher. Authors of submitted papers automatically accept a copyright transfer to *Chair of Production Engineering*, *University of Maribor*. For most up-to-date information on publishing procedure please see the *APEM journal* homepage *apem-journal.org*.

Submission of papers

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

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

Manuscript preparation

Manuscript should be prepared in *Microsoft Word 2007* (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 13 | Number 4 | December 2018 | pp 369-506

Contents

Scope and topics 372
Flexible job shop scheduling using zero-suppressed binary decision diagrams 373
Meolic, R.; Brezočnik, Z.
Functional objectives decision-making of discrete manufacturing system based on integrated ant colony optimization and particle swarm optimization approach 389
Xu, W.; Yin, Y.
Effect of process parameters on the surface roughness of aluminum alloy AA 6061-T6 sheets in frictional stir incremental forming 405
Azpen, Q.; Baharudin, H.; Sulaiman, S.; Mustapha, F.
Visual and optometric issues with smart glasses in Industry 4.0 working environment 417
Vujica Herzog, N.; Buchmeister, B.; Beharic, A.; Gajsek, B.
Multi-objective production planning model for equipment manufacturing enterprises with multiple uncertainties in demand 429
Liu, Y.F.; Zhang, Q.S.
A quantitative analysis method of greenhouse gas emission for mechanical product remanufacturing based on Petri net 442
Shi, J.L.; Fan, S.J.; Wang, Y.J.; Cheng, J.S.
Comprehensive analysis and study of the machinability of a high strength aluminum alloy (EN AW-AlZn5.5MgCu) in the high-feed milling 455
Duplák, J.; Hatala, M.; Dupláková, D.; Steranka, J.
Hybrid fruit fly optimization algorithm for solving multi-compartment vehicle routing problem in intelligent logistics 466
Wang, C.L.; Li, S.W.
Multi-objective transport network design with a reversible simulated annealing algorithm 479
Feng, X.; Ruan, Z.; Zhu, X.; Zhang, L.
Compatibility of ionic liquids with hydraulic system components 492
Kambič, M.; Kalb, R.; Tič, V.; Lovrec, D.
Calendar of events 504
Notes for contributors 505

ISSN 1854-6250

apem-journal.org