Swarm and Evolutionary Computation 69 (2022) 100985

Contents lists available at ScienceDirect

Swarm and Evolutionary Computation

journal homepage: www.elsevier.com/locate/swevo

# Multi-objective Q-learning-based hyper-heuristic with Bi-criteria selection for energy-aware mixed shop scheduling

Lixin Cheng [a,b], Qiuhua Tang [a,b,*], Liping Zhang [a,b], Zikai Zhang [a,b]

[a] Key Laboratory of Metallurgical Equipment and Control Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan, China
[b] Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan, China

ARTICLE INFO

*Keywords:*
Mixed shop
Energy-aware scheduling
Bi-criteria selection
Hyper-heuristic
Q-learning

ABSTRACT

Owning to diverse customer demands and enormous product varieties, mixed shop production systems are applied in practice to improve responsiveness and productivity along with energy-saving. This work addresses a mixture of job-shop and flow-shop production scheduling problem with a speed-scaling policy and no-idle time strategy. A mixed-integer linear programming model is formulated to determine the speed level of operations and the sequence of job-shop and flow-shop products, aiming at the simultaneous optimization of production efficiency and energy consumption. Then, a multi-objective Q-learning-based hyper-heuristic with Bi-criteria selection (QHH-BS) is developed to obtain a set of high-quality Pareto frontier solutions. In this algorithm, a new three-layer encoding is designed to represent the production sequence of job-shop and flow-shop products; the Pareto-based and indicator-based selection criteria are sequentially implemented to encourage diversity and convergence; Q-learning with a multi-objective metric-based reward mechanism is applied to select an optimizer from three prominent optimizers in each iteration for better exploration and exploitation. Three conclusions are drawn from extensive experiments: (1) Bi-criteria selection is superior to single-criterion selections; (2) Q-learning-based hyper-heuristic is more effective and robust than single optimizer-based algorithms and simple hyper-heuristics; (3) QHH-BS outperforms the existing state-of-the-art multi-objective algorithms in convergence and diversity.

## 1. Introduction

Flow-shop and job-shop are two types of production systems widely used in manufacturing systems. Flow-shop is useful for mass/flow-shop production which is characteristic with a fixed product flow and a relatively large quantity of production. Job-shop is suitable for one-of-a-kind production where each product is produced with a specific processing route. In the context of industry 4.0, many factories prefer to concurrently produce a mass of products with similar processing features and some personalized products to satisfy the diversified demands of customers [1]. The desired production system should process both flow-shop and job-shop products, and hence a mixed shop of flow-shop and job-shop is in need. In nature, as long as factories equip necessary transportation facilities and machines with higher generality, this kind of mixed shop can be implemented [2]. Although mixed shops have more applications in real production, studies on this kind of mixed shop scheduling problem are still in the infant stage to our knowledge.

Under the pressure of environmental issues and energy costs, energy saving attracts the attention of enterprises and researchers. Since manufacturing processes account for 33% of total energy consumption, energy-aware scheduling has emerged as an indispensable measure to save energy [3]. Motivated by this, abundant studies have been conducted on energy-aware job-shop or flow-shop scheduling problems to reduce energy consumption. Nevertheless, rare studies have been reported on solving the energy-aware mixed shop scheduling problem of flow-shop and job-shop. To fill this gap, this work addresses this emerging problem in which flow-shop and job-shop products coexist, for the purpose of improving productivity and saving energy.

To achieve the trade-off between productivity and energy consumption, an efficient multi-objective evolutionary algorithm (MOEA) is needed. Nevertheless, it is hard for most existing MOEAs to balance convergence and diversity. Pareto-based MOEAs maintain good diversity but suffer from slow convergence, while Non-Pareto-based MOEAs provide high convergence but may be trapped into local optimum [4]. Hence, Bi-criteria-based MOEA, which combines the advantages of Pareto-based MOEAs and Non-Pareto-based MOEAs, is expected to obtain a set of well-converged and well-distributed solutions. Besides, the performance of most MOEAs is highly dependent on the problem domain and heuristic techniques. If facing the problem instances from a variety of domains, they may lose their superiority [5]. In contrast,

* Corresponding author at: Key Laboratory of Metallurgical Equipment and Control Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan, China
*E-mail address:* tangqiuhua@wust.edu.cn (Q. Tang).

https://doi.org/10.1016/j.swevo.2021.100985
Received 23 April 2021; Received in revised form 17 July 2021; Accepted 11 September 2021
Available online 16 September 2021
2210-6502/© 2021 Published by Elsevier B.V.

hyper-heuristics are more applicable and it shows superiority in solving optimization problems by combining the advantages of different low-level optimizers. To this end, a multi-objective Q-learning-based hyper-heuristic with Bi-criteria selection is designed to solve the emerging energy-aware mixed shop scheduling problem. Against previous research, the contributions are summarized below.

(1) A mixed-integer linear programming model (MILP) is formulated for an energy-aware mixed shop scheduling problem with a speed-scaling policy and no-idle time strategy. In this model, three types of variables are involved: the global sequence for job-shop products, the local sequence for flow-shop products, and the speed level of operations.
(2) A multi-objective Q-learning-based hyper-heuristic with Bi-criteria selection (QHH-BS) is developed to solve the problem. In QHH-BS, Pareto-based and indicator-based criteria are sequentially executed in the selection stage to encourage diversity and convergence; Q-learning with a multi-objective metric-based reward mechanism is applied to select an appropriate optimizer in each iteration for better exploration and exploitation.

The rest of the study is organized as follows. Section 2 is the literature review. Section 3 describes the problem and develops a MILP model. Section 4 details the proposed QHH-BS. Section 5 reports experimental results. Section 6 draws a conclusion and presents future work.

## 2. Literature review

Till now, rare studies have been reported on energy-aware mixed shop scheduling problems. Hence, this section mainly reviews the mixed shop scheduling problems, the energy-saving strategies applied in shop scheduling problems and multi-objective optimization approaches.

In general, three types of production systems, open-shop, flow-shop and job-shop, are utilized to produce products with comprehensive processing routes. However, real-world production scenarios are usually more complex and tend to be a mixture of the above three shops [6]. Three mixed shops occur, namely the mixture of open-shop and flow-shop [7], the mixture of open-shop and job-shop [8] and the mixture of job-shop and flow-shop. Concerning the mixture of job-shop and flow-shop, most research regards the flow-shop as a special case of job-shop, and treats each flow-shop product as a job-shop one and hence transform this mixture into a simple job-shop [6]. However, when some realistic constraints are involved, the flow-shop behaves quite differently from the job-shop. Take the no-idle time constraint as an example. It requires the involved machines to work continuously to ensure production continuity and reduce setup times. This constraint is commonly utilized in flow-shop, which attributes to the similar processing features among the work-in-process products. Nevertheless, it must be stressed that this constraint is normally not adopted in job-shop. Hence, the job-shop and flow-shop products need to be treated differently, and the resulted shop is regarded as the mixture of job-shop and flow-shop. For this problem, Li and Wang [2] firstly addressed the mixture of no-idle flow-shop and job-shop production problems according to the characteristics of the real factory. However, this research focused more on system reconfiguration and production costs rather than production scheduling. Besides, the authors assumed that each flow-shop product has the same processing time, which limits the application of the model. For this reason, this work extends the model of [2] from the perspective of production scheduling, and further, energy consumption is considered to achieve green manufacturing.

Energy-saving can be achieved from the shop floor level and the machine level. Regarding the shop floor level, low-power processing machines can be selected from alternative machines to save processing energy, and the processing sequence of products can be optimized to reduce idle times of machines and thus to cut down standby energy consumption. Besides, the path of vehicles can be well planned to decrease transportation energy consumption [9]. From the machine level, the machine off/on criterion and the speed-scaling policy are two widely used energy-saving strategies. The machine off/on criterion works based on the fact that idle machines consume standby energy during idle times, and hence switching off/on the idle machines at the right time can save energy [10]. However, frequently turning on/off seriously shorten the life cycle of machines [11]. And this criterion is inapplicable for some production scenarios, i.e. when machines are not permitted to shut down. In this situation, the speed-scaling policy is considered and successfully applied in job-shop, flow-shop and parallel machine scheduling problems to save energy. Wu and Che [12] designed a speed adjusting heuristic for the energy-aware parallel machine scheduling problem. Chen and Wang [13] presented a non-critical operation-based speed adjusting strategy for the no-idle flow-shop problem. Abedi and Chiong [14] determined the speed level in an energy-aware job-shop scheduling problem. From the above literature, it is evident that the speed-scaling policy is efficient for saving energy in the studied energy-aware shop scheduling problems. However, rare studies focus on energy-aware mixed shop scheduling problems of job-shop and flow-shop. Hence, this work addresses the mixture of job-shop and no-idle flow-shop scheduling problem with the speed-scaling policy to minimize makespan and energy consumption simultaneously.

To solve multi-objective optimization problems efficiently, diverse MOEAs have been developed. According to the selection criterion, MOEAs are broadly classified as Pareto-based MOEAs and non-Pareto-based MOEAs. Pareto-based MOEAs, such as the non-dominated sorted genetic algorithm II (NSGA-II) [15], select the solutions via Pareto dominance rank and crowding distance. Non-Pareto-based MOEAs consist of indicator-based MOEAs and decomposition-based MOEAs. The indicator-based MOEAs, such as R2-MOEAs [16], construct a single performance indicator as the fitness function to evaluate and select solutions. The decomposition-based MOEAs, such as MOEA/D [17], decompose the multi-objective optimization problem into a set of single-objective subproblems and optimize them simultaneously. However, the above MOEAs are still far away from satisfactory in finding a group of well-converged and well-distributed solutions [18]. Pareto-based MOEAs maintain good diversity but suffer from slow convergence while non-Pareto-based MOEAs provide high convergence but may be trapped into local optimum [4]. In this situation, Bi-criteria-based MOEAs which combine Pareto-based MOEAs and non-Pareto-based MOEAs become active in recent years. In this area, Wang and Jiao [19] divided and stored the solution set in two archives. The convergence archive collected the solutions according to the $I_{\varepsilon+}$ indicator, and the diversity archive collected and updated solutions by Pareto dominance. Li and Yang [18] developed a novel Bi-criteria-based MOEAs, in which two populations shared and exchanged information frequently but evolved based on the Pareto-based criterion and indicator-based criterion, respectively. Cai and Li [20] proposed a hybrid MOEA, in which a decomposition-based method was adopted for the working population and a Pareto-based mechanism was applied to maintain the external archive. In these Bi-criteria-based MOEAs, two selection criteria executed simultaneously in each iteration, and hence more computing resources were required for selection. Thus, a modified Bi-criteria-based MOEA is in need in this work for a better compromise between diversity and convergence under limited computing resources.

Another problem in the existing MOEAs is that their performance is highly dependent on the problem domain and heuristic techniques [21,22]. If facing problem instances from a variety of domains, they may lose their superiority [5]. In this context, hyper-heuristics are more generally applicable since they require no problem-specific knowledge. In general, hyper-heuristics consist of a group of low-level heuristics. For any given problem, the core task of hyper-heuristics is to determine the performing sequence of the low-level heuristics rather than to solve the problem directly. The communication between the low-level heuristic and the hyper-heuristic is achieved through a problem-independent interface, which separates the problem domain and the

hyper-heuristic, and raises the level of generality [23]. Moreover, each low-level heuristic behaves differently in exploration and exploitation, hyper-heuristics benefit from the advantages of different heuristics and guide the search process into the promising regions of different problems. Therefore hyper-heuristics are more effective than the general MOEAs. Nevertheless, most hyper-heuristics, such as the backtracking search-based hyper-heuristics [24] and the genetic programming-based hyper-heuristics [25], searched an appropriate performing sequence of heuristics via a trial-and-error method. The search process of these hyper-heuristics is time-consuming and the experience obtained from the search process cannot be accumulated and used for self-learning. Recently, reinforcement learning has been successfully applied to hyper-heuristics to intelligently and adaptively select low-level heuristics. In this area, Qu and Gai [26] developed a reinforcement learning-based grey wolf optimizer, which included four operations, exploration, exploitation, geometric adjustment and optimal adjustment, and employed reinforcement learning to select an appropriate operation. Glcük and Ozsoydan [27] utilized four state-of-the-art bio-inspired metaheuristics as the low-level optimizers and applied Q-learning to select an optimizer in each iteration. Motivated by these successful applications of hyper-heuristics, a modified multi-objective Q-learning-based hyper-heuristic algorithm is designed in this work to solve this new energy-aware mixed shop scheduling problem with two expected advantages. (1) The existing operators, which have been proved to be effective in solving similar energy-aware scheduling problems, can be utilized directly as low-level optimizers. (2) A problem-independent hyper-heuristic, which determines the performing sequence of low-level heuristics, is robust enough and easy to implement.

The above review reveals that the mixture of job-shop and no-idle flow-shop scheduling problems receives little attention although it has important applications in real industries. Second, energy-aware scheduling becomes a research hotspot in recent years, but rare studies focus on energy-aware mixed shop scheduling of job-shop and flow-shop. Third, it is imperative to integrate the existing optimizers, which solved similar energy-aware scheduling problems successfully, into a hyper-heuristic solution framework so as to solve the new energy-aware mixed shop scheduling problem efficiently. Therefore, a multi-objective Q-learning-based hyper-heuristic with Bi-criteria selection is tailored to solve the considered energy-aware mixed shop scheduling problem in which speed-scaling policy and no-idle time strategy are both adopted.

## 3. Problem description and formulation

### 3.1. Problem description

In a mixed shop, two types of products, job-shop and flow-shop products, are processed. Each job-shop product has several operations and needs to be processed in a specific processing route, which is illustrated by the four job-shop products $\{j_1, j_2, j_3, j_4\}$ in Fig. 1. In addition, all flow-shop products possess the same processing route and are required to be processed sequentially on machines in the same order to reduce the setup times among products. More important, idle time between two successive flow-shop products is not allowed. The purpose of no-idle time restriction is to improve the operational continuity of machines and reduce the number of machine shutdowns.

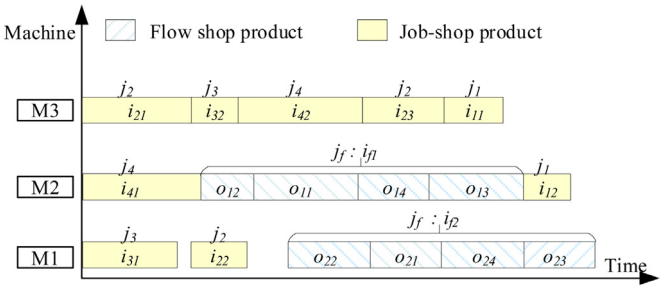


**Fig. 1.** Mixed shop scheduling involving job-shop and flow-shop products.

This mixed shop scheduling problem requires determining the processing sequence of two types of products. To simplify this problem, all flow-shop products processed continuously are viewed as a batch and hence a batch of flow-shop products is treated as a job-shop product. As a result, all job-shop and flow-shop products can be scheduled globally as in the general job-shop scheduling problem. As shown in Fig. 1, there are five job-shop products, $\{j_1, j_2, j_3, j_4, j_f\}$ where $j_f$ denotes a batch of flow-shop products. Note that in this batch, all the flow-shop products should be sequenced locally and the no-idle time constraint should be satisfied to ensure the operational continuity of machines. Hence, the mixed shop scheduling problem of job-shop and flow-shop is divided into two subproblems: globally scheduling of all job-shop products and locally sequencing of flow-shop products.

In addition, each machine is speed-scalable and has several alternative speed levels. A higher speed may reduce the processing time but consume more energy. Hence, the determination of speed level for each operation plays a determinant role in the energy efficiency within this mixed shop. In a word, this work aims to minimize the total energy consumption and makespan simultaneously by scheduling globally all job-shop products, sequencing locally the flow-shop products and determining the speed level of each operation.

### 3.2. Notations

Sets:

| | |
|---|---|
| $\mathbf{J}$ | Set of products, $\mathbf{J}=\{1, 2,\ldots j\ldots, J, f\}$. It contains job-shop products ($1 \le j \le J$) and a batch of flow-shop products ($j=f$). |
| $\mathbf{I_j}$ | Set of operations of product $j$, $\mathbf{I_j}=\{1, 2,\ldots i_j\ldots, I_j\}$. |
| $\mathbf{M}$ | Set of machines, $\mathbf{M}=\{1, 2,\ldots m\ldots, M\}$. |
| $\mathbf{Q}$ | Set of products within a batch, $\mathbf{Q}=\{1, 2,\ldots q\ldots, Q\}$. |
| $\mathbf{K}$ | Set of speed levels, $\mathbf{K}=\{1, 2,\ldots k\ldots, K\}$. |

Parameters:

| | |
|---|---|
| $\alpha$ | The exponent of processing energy consumption. |
| $n$ | The number of flow-shop products within a batch, $n = \lvert\mathbf{Q}\rvert$. |
| $p_{jim}$ | Normal processing time of operation $i_j$ on machine $m$. |
| $h_{iqm}$ | Normal processing time of the $i$-th operation of the $q$-th flow-shop product on machine $m$, and $p_{jim} = \sum_{q=1}^{n} h_{iqm},\ if\ j = f$. |
| $w_m$ | Normal processing power by machine $m$. |
| $v_k$ | The processing speed at speed level $k$ on a machine, $v_k \in \{v_1, v_2, \ldots v_K\}$ |
| $\varphi$ | A large enough positive number. |

Decision variables:

| | |
|---|---|
| $S_{mjij'i'}$ | Take value 1 if operation $i_j$ immediately precedes operation $i'_{j'}$ on machine $m$; otherwise, 0. |
| $Z_{qt}$ | Take value 1 if the $q$-th flow-shop product occupies position $t$ in the order of flow-shop products; otherwise, 0. |
| $U_{jik}$ | Take value 1 if operation $i_j$ is processed on a machine at speed level $k$. |

Intermediate variables:

| | |
|---|---|
| $T^s_{ji}$ | Start time of operation $i_j$. |
| $T^e_{ji}$ | Finish time of operation $i_j$. |
| $T^b_{iq}$ | Start time of the $i$-th operation of the $q$-th flow-shop product. |

### 3.3. Model formulation and linearization

Based on the above notations, a mathematical model of energy-aware mixed shop scheduling problem with the speed-scaling policy and no-idle time strategy is formulated as follows. Note that, according to the speed-scaling policy, if a product is processed on a machine at speed

level $v_k$, the actual processing time of the product is $p_{jim}/v_k$, and the corresponding processing power is $w_m \cdot (v_k)^{\alpha}$ [28].

$$min\ Cmax = max_{j,i|(i\in I_j)} T^e_{ji} \tag{1}$$

$$\min TEC = \sum_{m=1}^{|M|}\sum_{k=1}^{|K|}\sum_{j=1}^{|J|}\sum_{i=1}^{|I_j|}\frac{P_{jim}}{v_k}\cdot U_{jik}\cdot w_m\cdot (v_k)^{\alpha} \tag{2}$$

***s.t.***

$$\sum_{k=1}^{|K|} U_{jik} = 1, \quad \forall j \in J,\ i \in I_j \tag{3}$$

$$S_{mjij',i'} + S_{mj',i',ji} = 1, \forall\ (j,j') \in J, i \in I_j,\ i' \in I_{j'},\ (j \neq j'\ \ or\ (j = j'\ and\ i < i')), m \in M_{ji},\ m' \in M_{j',i'} \tag{4}$$

$$T^s_{j',i'} \geq T^e_{ji} - \varphi\cdot(1 - S_{mji,j',i'})\forall(j,j') \in J, i \in I_j,\ i' \in I_{j'},\ (j \neq j'\ \ or\ (j = j'\ and\ i < i')) \tag{5}$$

$$T^e_{ji} = T^s_{ji} + \sum_{k=1}^{|K|} U_{jik}\cdot\frac{p_{jim}}{v_k}, \quad \forall j \in J,\ i \in I_j \tag{6}$$

$$T^s_{j,i+1} \geq T^e_{ji}, \quad \forall(1 \leq j \leq J),\ (i,i+1) \in I_j,\ i < \left|I_j\right| \tag{7}$$

$$T^b_{i+1,\,q} \geq T^b_{iq} + \sum_{k=1}^{|K|}\frac{h_{iqm}}{v_k}\cdot U_{jik}, \quad \forall j = f,\ i \in I_j,\ q \leq n,\ i < \left|I_j\right| \tag{8}$$

$$\sum_{q=1}^{n} Z_{qt} = 1, \quad \forall\ t \leq n \tag{9}$$

$$\sum_{t=1}^{n} Z_{qt} = 1, \quad \forall\ q \leq n \tag{10}$$

$$T^b_{iq} \geq T^s_{ji} + \sum_{m=1}^{|M|}\sum_{k=1}^{|K|}\sum_{t'=1}^{t-1}\sum_{q'=1}^{n}\frac{h_{iq'm}}{v_k}\cdot U_{jik}\cdot Z_{q't'} - \varphi\cdot(1 - Z_{qt}), \quad \forall j = f,\ i \in I_j, q \leq n \tag{11}$$

$$T^b_{iq} \leq T^s_{ji} + \sum_{m=1}^{|M|}\sum_{k=1}^{|K|}\sum_{t'=1}^{t-1}\sum_{q'=1}^{n}\frac{h_{iq'm}}{v_k}\cdot U_{jik}\cdot Z_{q't'} + \varphi\cdot(1 - Z_{qt}), \quad \forall j = f,\ i \in I_j, q \leq n \tag{12}$$

The first objective in Eq.(1) is to minimize the makespan, and the second objective in Eq.(2) aims to minimize the total energy consumption. Constraint (3) assures that each operation is processed by a machine at a certain speed level. Constraints (4-5) arrange the sequence of two operations to be processed on the same machine. It demands that only one operation is processed by each machine at a time, and if operation $i_j$ precedes $i'_{j'}$ on a machine, the operation $i'_{j'}$ can start only after the operation $i_j$ has finished. Constraint (6) defines that the finish time of each operation equals its start time plus the processing time. Constraints (7-8) emphasize that an operation of job-shop products or flow-shop products cannot start until its previous operation is finished. Constraints (9-10) determine the sequence of flow-shop products. Constraints (11-12) enforce that idle time is not permitted between two successive flow-shop products, and in a flow-shop batch, the start time of position ($t$+1) strictly equals the completion time of position $t$.

The above model is not linear due to the item contained in constraints (11-12), $Z_{qt}\cdot U_{jik}$. Due to this, a variable $C_{jiqtk}$ is introduced as shown in constraints (13-14) to replace the multiplication of $Z_{qt}$ and $U_{jik}$, and constraints (15-17) limit $C_{jiqtk}$ take the value zero or one. Thereby, constraints (11-12) are completely replaced by constraints (13-17) and the nonlinearity disappears.

$$T^b_{iq} \geq T^s_{ji} + \sum_{m=1}^{|M|}\sum_{k=1}^{|K|}\sum_{t'=1}^{t-1}\sum_{q'=1}^{n}\frac{h_{iq'm}}{v_k}\cdot C_{jiq',t'k} - \varphi\cdot(1 - Z_{qt}), \quad \forall j = f,\ i \in I_j, q \leq n \tag{13}$$

$$T^b_{iq} \leq T^s_{ji} + \sum_{m=1}^{|M|}\sum_{k=1}^{|K|}\sum_{t'=1}^{t-1}\sum_{q'=1}^{n}\frac{h_{iq'm}}{v_k}\cdot C_{jiq',t'k} + \varphi\cdot(1 - Z_{qt}), \quad \forall j = f,\ i \in I_j, q \leq n \tag{14}$$

$$C_{jiqtk} \leq Z_{qt}, \quad \forall j = f, i \in I_j, k \in K, (t,q) \leq n \tag{15}$$

$$C_{jiqtk} \leq U_{jik}, \quad \forall j = f, i \in I_j, k \in K, (t,q) \leq n \tag{16}$$

$$C_{jiqtk} \geq Z_{qt} + U_{jik} - 1, \quad \forall j = f, i \in I_j, k \in K, (t,q) \leq n \tag{17}$$

Overall, with constraints (1-10, 13-17), the considered problem is formulated as a mixed-integer linear programming model (MILP).

## 4. Multi-objective Q-learning-based hyper-heuristic with Bi-criteria selection

To obtain a set of Pareto frontier solutions in a timely manner, a multi-objective Q-learning-based hyper-heuristic with Bi-criteria selection (QHH-BS) is proposed. The novelty of QHH-BS lies in three aspects: the problem-oriented three-layer encoding and decoding, the integration of Q-learning and multiple optimizers in the reproduction step, and the utilization of Bi-criteria in the selection step. Details of QHH-BS are introduced in the following sections.

### 4.1. Three-layer encoding

For the considered problem, continuous encoding is adopted for its validity and universal applicability in solving combinatorial optimization problems [29]. Three decision variables in the MILP model are encoded into three vectors respectively, namely global product sequence vector, local product sequence vector and speed level vector.

The global product sequence vector sequences all job-shop products and a batch of flow-shop products. In this vector, a batch of flow-shop products works like one job-shop product. This vector is encoded with $\sum_j I_j$ elements and each element is uniformly generated in the interval [0, $mn$], where $mn$ represents the largest number of operations among all products. By sorting these elements in the non-increasing order, the processing sequence of operations can be determined. Note that, an operation of a job can start after its previous one has finished, hence the randomly generated element for this operation must be less than that of the previous one.

The local product sequence vector is responsible for sequencing all flow-shop products within a batch. It consists of $n$ elements, where $n$ is the number of flow-shop products. Each element is generated from a uniform distribution [0, $n$]. The processing priority of each product is determined by sorting this vector in the non-increasing order.

The speed level vector determines the processing speed of machines. It consists of $\sum_j I_j$ elements and each element is drawn out of a uniform distribution [1, $k$+1], where $k$ represents the number of speed levels. After rounding down this vector, the resulted value is employed as the processing speed level of the corresponding operation.

To illustrate this encoding clearly, a case with one job-shop product (Job 1) and one batch of flow-shop products (Job $f$) are presented in Fig. 2. In the batch, there are three flow-shop products. Each job-shop or flow-shop product has two operations ($i_{11}$, $i_{12}$; $i_{f1}$, $i_{f2}$) and each operation must be performed at one of three alternative speed levels ($v_1$, $v_2$, $v_3$) to save energy. As shown in Fig. 2, the global product sequence vector, (1.8, 0.1, 1.4, 0.7), represents the priorities of performing the four operations. By sorting these priorities, the final sequence of these four operations is ($i_{11}$, $i_{f1}$, $i_{f2}$, $i_{12}$). Similarly, the final sequence of three flow-shop products is ($q_3$, $q_1$, $q_2$). For the speed level vector, each continuous number should be rounded down and hence the final speed levels for corresponding operations are ($v_1$, $v_2$, $v_1$, $v_3$).

### 4.2. Forward and backward decoding

The decoding process is decomposed into forward and backward decoding. The forward decoding ignores the no-idle time restriction and determines the earliest start time for each product under the condition that the assigned machine is available and the product has been released

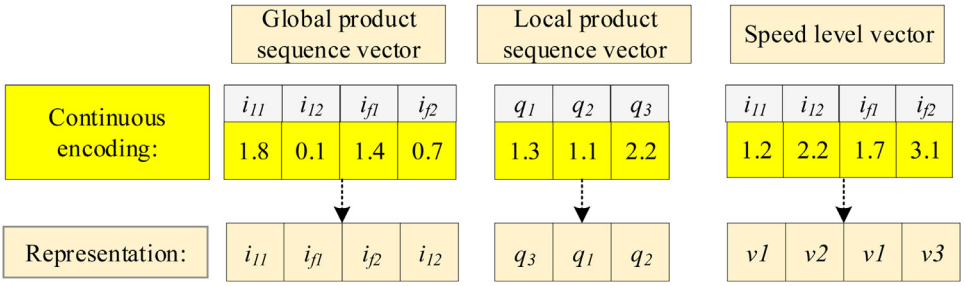


**Fig. 2.** An illustrative example of encoding.

from its previous operation. Subsequently, the backward decoding pulls backward the above start times in a batch from the last product to the first to meet the no-idle time constraint. The concrete steps of the decoding mechanism are summarized below. Note that steps 1-4 are forward decoding and step 5 means backward decoding.

**Step 1**: Take an operation from the global sequence vector and its speed level from the speed level vector.
**Step 2**: Calculate the actual processing time and the processing energy consumption of this operation.
**Step 3**: Calculate the earliest start time of this operation according to the completion time of its previous operation and the available time of the machine that it is assigned to.
**Step 4**: If the product is a job-shop product, go to step 6. Otherwise, check whether the start times of all flow-shop products in the local sequence vector are determined. If yes, go to step 5; otherwise, go to step 3.
**Step 5**: Fix the start time of the last product in the local sequence vector. Pull backward its previous one to ensure that no idle time exists between them. Repeat the process of pulling backward till the first product in the local sequence vector.
**Step 6**: If all products in the global sequence vector are scheduled, calculate the total energy consumption and makespan according to Eqs. (1-2); otherwise, go back to step 1.

### 4.3. Low-level optimizers with different search patterns

Recently, grey wolf operator, Jaya operator and genetic operator have been reported to successfully solve the energy-aware scheduling problems in flow shop [30], flexible job-shop [31] and job-shop [32] respectively. In particular, the grey wolf operator makes full use of the three best solutions and shows superiority in solution quality and convergence speed [33]; the Jaya operator is parameter-less and makes the most of the best and worst solutions; the crossover operation in genetic operator has strong exploitation ability, and works by exchanging information pieces between two randomly selected parents. Different optimizers perform differently in the search process, and hence these three optimizers are all incorporated into the proposed QHH-BS to solve the considered problem. In the following, the search patterns of these three optimizers are demonstrated briefly.

In grey wolf operator, three special individuals including the best individual $\vec{X}_\alpha$, the second-best individual $\vec{X}_\beta$ and the third-best individual $\vec{X}_\delta$, play a leading role in reproduction. Each individual $\vec{X}$ is updated toward the promising areas according to Eqs. (18-22).

$$\vec{X}_{i+1} = \left(\vec{X}_1 + \vec{X}_2 + \vec{X}_3\right)/3 \tag{18}$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot \vec{D}_\alpha,\ \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \vec{D}_\beta,\ \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot \vec{D}_\delta \tag{19}$$

$$\vec{D}_\alpha = \left|\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}\right|,\ \vec{D}_\beta = \left|\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}\right|,\ \vec{D}_\delta = \left|\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}\right| \tag{20}$$

$$\vec{A}_k = 2 \cdot \vec{a} \cdot \vec{r}_1 - \vec{a}, \quad k = 1,\ 2,\ 3 \tag{21}$$

$$\vec{C}_k = 2 \cdot \vec{r}_2, \quad k = 1,\ 2,\ 3 \tag{22}$$

where $\vec{D}_\alpha$, $\vec{D}_\beta$ *and* $\vec{D}_\delta$ denote the distance between a common solution and the three best solutions. $\vec{a}$ decreases linearly from 2 to 0 during the process of evolution, both $\vec{r}_1$ and $\vec{r}_2$ take random values from the interval [0, 1].

Jaya operator utilizes the best and worst solutions to update the population. As described in Eq.(23), it demands that each individual should move towards the best one and away from the worst one.

$$\vec{X}_{i+1} = \vec{X}_i + \vec{r}_1 \cdot \left(\vec{X}_b - \left|\vec{X}_i\right|\right) - \vec{r}_2 \cdot \left(\vec{X}_w - \left|\vec{X}_i\right|\right) \tag{23}$$

where $\vec{X}_b$ and $\vec{X}_w$ represent the best and worst solutions of the current population, $\vec{r}_1$ and $\vec{r}_2$ take random values in the interval [0, 1].

Crossover operator begins with a random selection of two individuals from the parent population, and then randomly swaps some fragments of the two individuals by a two-point crossover [34].

### 4.4. Q-learning-based hyper-heuristic

Instead of blindly stacking multiple optimizers, Q-learning intelligently recommends the most appropriate optimizer in each evolutionary state of the algorithm to achieve a better exploration and exploitation. In short, a Q-learning-based hyper-heuristic is studied in this work. In Q-learning, an intelligent agent perceives the state $s_t$ of the algorithm and takes an action $a_t$ from an action set $\boldsymbol{A}$ according to $\epsilon-$greedy policy. After the action $a_t$ is executed, the algorithm goes into a new state $s_{t+1}$, and returns a reward $r_t$ according to a well-designed reward mechanism. Meanwhile, a Q-table is updated to record the Q-value $Q_t(s_t, a_t)$ of each action $a_t$ in state $s_t$ as Eq.(24).

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \lambda \cdot \left[r_t + \gamma \max_a Q_t(s_{t+1}, a) - Q_t(s_t, a_t)\right] \tag{24}$$

where $\lambda$ and $\gamma$ represent the learning rate and the discount factor respectively, and all the initial Q-values are set as zero for fairness.

The goal of Q-learning is to determine a Q-table, and the agent chooses the action with the largest Q-values in the Q-table to maximize the cumulative rewards. To train the Q-table for the considered problem, three key elements of Q-learning, including action, state and reward mechanism, are designed elaborately and illustrated as follows.

The action set includes three optimizers, {grey wolf operator, Jaya operator, crossover operator}. In each state, one action (an optimizer) is selected according to $\epsilon-$greedy policy to reproduce new solutions.

The evolution state of the algorithm is measured by Eq.(25), and it takes any continuous values within the interval [0, 1]. To reduce the complexity of training, the continuous interval [0, 1] is discretized into $k$ equally spaced subintervals. And hence, the state size is $k$ and the state set is represented as $\mathbf{S}=\{s_1, s_2, \ldots s_k\}$, where the interval value is $1/k$, and $s = s_1,\ if\ s^* \in [0,\ 1/k);\ s = s_k,\ if\ s^* \in ((k-1)/k,\ 1]$.

$$s^* = ct/st \tag{25}$$

where $ct$ is the cumulative running time and $st$ represents a given termination time.

After performing an action, new solutions may be generated and the Pareto front may be updated. Therefore, the Hypervolume Ratio [35] of the previous and new Pareto fronts are introduced to evaluate the action. A multi-objective metric-based reward mechanism is designed as

Eq.(26) to reward the action if the Pareto front is improved after performing the action; otherwise, to punish the action.

$$r_t = \begin{cases} \frac{HVR(P_{t+1})}{HVR(P_t)}, & if\ HVR(P_{t+1}) > HVR(P_t) \\ \frac{HVR(P_{t+1})}{HVR(P_t)} - 1, & otherwise \end{cases} \tag{26}$$

where $P_t$ and $P_{t+1}$ represent the Pareto front of the population in the $t$-th and the ($t$+1)-th iteration. $HVR(P_t)$ and $HVR(P_{t+1})$ are the corresponding Hypervolume Ratio. Note that $HVR$ is a positive value in [0, 1] and a larger HVR value indicates a better convergence and diversity.

### 4.5. Bi-criteria selection

In evolutionary algorithms, the selection operator picks out the solutions with better comprehensive performance into the next iteration. There are two commonly utilized selection criteria, the indicator-based and the Pareto-based. The Pareto-based criterion ranks solutions into different levels and can obtain solutions with higher diversity, while the indicator-based criterion prefers the solution with the better indicator value and hence can provide higher convergence pressure [4]. Inspired by this, Bi-criteria selection is proposed in this work to achieve faster convergence and higher diversity simultaneously. This strategy is implemented by dividing the evolutionary process into two stages. The earlier stage selects the excellent solutions into the next iteration by the Pareto-based criterion while the later stage adopts the indicator-based criterion.

For better understanding of the Pareto-based selection criterion, the non-dominated sorting mechanism is given in advance. The non-dominated sorting mechanism ranks the solutions into different levels according to the Pareto dominance relationship. It demands that the solutions at the lower-level Pareto dominate those at a higher level while the solutions at the same level are non-dominated. After all the solutions are assigned to different levels, the Pareto-based criterion prefers to select the solutions with lower levels; and if two solutions are at the same level, the one with the larger crowding distance is picked out [36].

The indicator-based criterion uses a single performance indicator to evaluate each solution. The commonly used indicators are R2-indicator [37] and HV-indicator [38]. R2-indicator assesses the contribution of each solution by a utility function. The weighted sum utility function approach is utilized in this work. To calculate the contribution of a solution $x$, the utility value of each solution $x$ in the solution set A is calculated by Eq.(27). And then, $R2(A,V)$ and $R2(A\backslash\{x\},V)$ of the solution set A and the solution set A\{x} are derived by Eq.(28). Finally, the contribution of solution $x$, $RC(x)$, is obtained by Eq.(29).

$$u(x|w) = \sum\nolimits_{1\le i\le m} w_i \cdot f_i(x) \tag{27}$$

$$R2(A,V) = \frac{1}{|V|} \cdot \sum\nolimits_{w\in V} \min\nolimits_{x\in A} u(x|w) \tag{28}$$

$$RC(x) = R2(A,V) - R2(A\backslash\{x\},V) \tag{29}$$

where $f_i(x)$ is the objective value of solution $x$ on objective $i$; $m$ is the number of involved objectives; $w$ denotes a randomly given weight vector, $w = (w_1, w_2, \ldots w_m)$ *and* $\sum_{i=1}^{m} w_i = 1$; $V$ is made up of a set of weight vectors $w$. The dimension of $V$, also called weight dimension, is determined through parameter calibration.

According to Eqs.(27-29), the contribution of each solution, $RC(x)$, takes a positive value or equals zero. The solutions with larger contributions have a higher priority to be selected. However, it is hard to select a solution from those with contributions equal to zero. Moreover, the computational cost increases significantly as the size of the solution set grows. To cope with the above limitations, an aggregation-based indicator is designed and defined as Eqs.(30-32). This indicator inherits the weighted sum utility function of the R2-indicator. It calculates the utility values of a solution under a set of weight vectors according to Eq.(31), and then takes the minimum utility value as the fitness of the solution. Those solutions with a smaller fitness have a higher priority to be selected.

$$min\_i = min f_i(\cdot), \quad max\_i = max f_i(\cdot) \tag{30}$$

$$u^*(x|w) = \sum\nolimits_{1\le i\le m} w_i \cdot \left( \frac{f_i(x) - min\_i}{max\_i - min\_i} \right) \tag{31}$$

$$fit(x) = \frac{1}{|V|} \min\nolimits_{w\in V} u^*(x|w) \tag{32}$$

In comparison, the computation complexity of Eqs.(27-29) is $2m$, $2m \cdot |A| \cdot |V|$ and $4m \cdot |A| \cdot |V|$ respectively while that in Eqs.(30-32) is $2|A|$, $4m|$ and $4m \cdot |V|$ respectively. Hence, the computation complexity of the R2-indicator-based criterion is $4m \cdot |A| \cdot |V|$ but that of the aggregation-based indicator is $2|A|+4|m| \cdot |V|$. Obviously, the aggregation-based indicator has a smaller computation complexity and hence is more suitable for the application.

In sum, Bi-criteria selection employs the Pareto-based criterion in the earlier search process for higher diversity while adopts the aggregation-based indicator in the later process for faster convergence and less computation.

### 4.6. Framework

This proposed Q-learning-based hyper-heuristic with Bi-criteria selection (QHH-BS) mainly includes two operations: reproducing offspring population with a recommended low-level optimizer and selecting solutions with better comprehensive performance. By repeating these two operations, a set of well-converged and well-distributed solutions can be achieved. Algorithm 1 shows the detailed framework of this algorithm.

## 5. Experimental study

To verify the effectiveness of QHH-BS for the energy-aware mixed shop scheduling problem with the speed-scaling policy and no-idle time strategy, five sets of experiments are performed. First, a real case is solved by QHH-BS to provide some insights into the considered problem. Second, the best parameter combinations of QHH-BS and the comparison MOEAs are calibrated. Third, the effect of the two improvements on QHH-BS is investigated separately. Subsequently, the validity of the proposed MILP model and QHH-BS is verified on small-scaled problems. Finally, QHH-BS is compared with other state-of-the-art MOEAs.

### 5.1. Experimental settings

Since this work addresses a mixture of job-shop and flow-shop production, testing cases marked as La#-F# are constructed by a combination of job-shop and flow-shop benchmark cases. Specifically, job-shop cases labeled as La01-La30 are taken from the sdata of Hurink_Data on the website http://www.idsia.ch/~monaldo/fjsp.html. Flow-shop cases marked as F1-F6 are taken from the website http://mistic.heig-vd.ch/taillard/problemes.dir/ordonnancement.dir/ordonnancement.html. Note that, F1 and F2 are the cases with 20 jobs and 5 machines but with different processing times. Similarly, F3 and F4 are the cases with 20 jobs and 10 machines; and F5 and F6 represent those with 50 jobs and 5 machines. In addition, the processing power $w_m$ is generated by a uniform distribution in [5,20] and the processing energy consumption exponent $\alpha$ takes the value of 1.2. The speed level set $v_k$ is set as {0.8, 1, 1.2}.

The proposed MILP model is solved by GAMS/Cplex win64 24.8.5 and all the other algorithms are coded in MATLAB R2018a and run on an Intel(R) Core(TM) i5 CPU processor with 2.80 GHz and 4.00GB RAM. Except for special instructions, the stopping criterion of all the comparison algorithms is set as $\sum_j I_j \times$ 200 ms.

**Algorithm 1**
QHH-BS.

01: Initialize parameters, evolutionary state, and population $P$ of size $N$.
02: Pick out the non-dominated solutions to construct the archive set.
03: **While** not satisfy the termination criterion **do**
// **Q-learning-based hyper-heuristic**
04: Recommend the best optimizer for the current state from Q-table.
05: **Switch** optimizers
06: **case 1**: Apply the grey wolf operator to $P$ to create offspring population $Q$.
07: **case 2**: Apply the Jaya operator to $P$ to create offspring population $Q$.
08: **case 3**: Apply the crossover operator to $P$ to create offspring population $Q$.
09: **End Switch**
10: If the offspring is beyond the given range, scale it to fit the interval. Otherwise, go to the next step.
11: Form a combined population $R$ of size $2N$, $R = P \cup Q$.
// **Bi-criteria selection**
12: **If** not reach *the mixing ratio* **do**
13: Select $N$ individuals from $R$ to create a new population $P$ via the Pareto-based criterion.
14: **else**
15: Select $N$ individuals from $R$ to create a new population $P$ via the aggregation-based indicator.
16: **End If**
17: Update the archive set by comparing population $P$ with the previous non-dominated solutions.
18: Update the state.
19: **End While**
20: **Return** archive set.

**Table 1**
Processing parameters of the real case.

| | Opera-tions | Jobs ina batch | Turn machine | Drill machine | Mill machine | Fine mill machine | Cutting machine |
|---|---|---|---|---|---|---|---|
| $w_m$ | | | 10 | 20 | 12 | 15 | 18 |
| $j_1$ | ($i_{11}$-$i_{15}$) | | $i_{12}$ (53) | $i_{11}$ (30) | $i_{14}$ (34) | $i_{15}$ (60) | $i_{13}$ (50) |
| $j_2$ | ($i_{21}$-$i_{25}$) | | $i_{21}$ (50) | $i_{22}$ (15) | $i_{23}$ (52) | $i_{24}$ (75) | $i_{25}$ (70) |
| $j_3$ | ($i_{31}$-$i_{35}$) | | $i_{34}$ (60) | $i_{33}$ (20) | $i_{31}$ (39) | $i_{35}$ (40) | $i_{32}$ (30) |
| $j_4$ | ($i_{41}$-$i_{45}$) | | $i_{43}$ (30) | $i_{45}$ (20) | $i_{42}$ (15) | $i_{44}$ (30) | $i_{41}$ (40) |
| $j_f$ | ($i_{f1}$ - $i_{f2}$) | ($q_1$-$q_6$) | | $o_{11}$(42); $o_{12}$(15); $o_{13}$(20); $o_{14}$(30); $o_{15}$(10); $o_{16}$(25) | | $o_{21}$(30); $o_{22}$(25); $o_{23}$(20); $o_{24}$(14); $o_{25}$(40); $o_{26}$(25) | |

To evaluate both the convergence and diversity of algorithms, the inverted generational distance (*IGD*) [39] and Hypervolume Ratio (*HVR*) [35] are adopted. For the two indicators, a smaller *IGD* value and a larger *HVR* value are expected. In addition, *C Metric* defined by Eq.(33) is employed to reveal the Pareto dominance relationship between the proposed algorithm and the comparison algorithms.

$$C(A, B) = |\{a \in A; \exists b \in B : a \succ b\}|/|B| \tag{33}$$

where *C(A, B)* takes value in [0, 1]. It denotes the proportion of solutions in $B$ that are dominated by solutions from $A$. The more solutions in $B$ are dominated by the solutions in $A$, the closer the value of *C(A, B)* is to 1. If *C(A, B)* equals 1, set B is fully dominated by set A and hence set B is inferior to set A.

### 5.2. A real-world case study

To gain a better insight into the considered problem, a small case extracted from the real-world production is demonstrated and solved by QHH-BS. In this case, four job-shop products (named $j_1$-$j_4$) are processed on five machines according to specific process routes as represented in Table 1. Take $j_1$ as an example, its process route is Drill→Turn→Cutting→Mill→Fine mill. Additionally, there are six flow-shop products (named $q_1$-$q_6$) and each of them is processed by two machines according to a fixed process route. Since the six products have similar processing requirements, they are viewed as a small batch and named as $j_f$. The normal processing time (*min*) of each operation and processing power (*kw*) of each machine are described in Table 1. Each machine has three speed levels (*v1-v3*), and *v1, v2* and *v3* take 1, 1.2 and 0.8. The processing energy consumption exponent $\alpha$ is 1.2.

The obtained three reprehensive non-dominated solutions are {A*: *Cmax*=345, *TEC*=17044; B*: *Cmax*=349.2, *TEC*=16989; C*: *Cmax*=351.7, *TEC*=16767}. Clearly, solution A*/C* has the smallest *Cmax/TEC* respectively. Figs.(3-4) show the Gantt charts of solutions A* and C*. In Fig. 3, all operations on the critical path are processed with the highest speed, which leads to a smaller *Cmax* but a higher *TEC*. Instead, in Fig. 4, some operations on the critical path are processed with a relatively lower speed, and most of the other operations are processed with the lowest speed. As a result, at the cost of a smaller increase of *Cmax, TEC* in Fig. 4 is far less than that in Fig. 3.

It can be concluded that when more operations are processed on the machines with lower speed, the obtained production plan consumes less energy, but results in a longer makespan due to the increase of the corresponding processing times of these operations. In contrast, if the machines on the critical path are running at a higher speed, the obtained production plan has a smaller makespan at the cost of larger energy consumption. Therefore, it is imperative to develop some hyper-heuristics so that a delicate balance between makespan and energy consumption can be achieved even if the problem domain changes.

### 5.3. Parameter calibration of QHH-BS and the comparison MOEAs

QHH-BS includes seven parameters. The population size is the basic parameter; the mixing ratio and weight dimension come from the Bi-criteria selection; the state size, learning rate, discount factor and greedy factor are critical for the training of Q-learning. To find the best parameter combination of these parameters, the Taguchi method of design of experiments [40] is employed. Each parameter is set to have three levels and the values of all levels are reported in Table 2.

QHH-BS is conducted on six cases {La30-F1, La30-F2, La30-F3, La30-F4, La30-F5, La30-F6}, and each case runs ten times for each parameter combination to obtain an average performance. The average *IGD* and *HVR* are taken as the response variables and recorded in the last

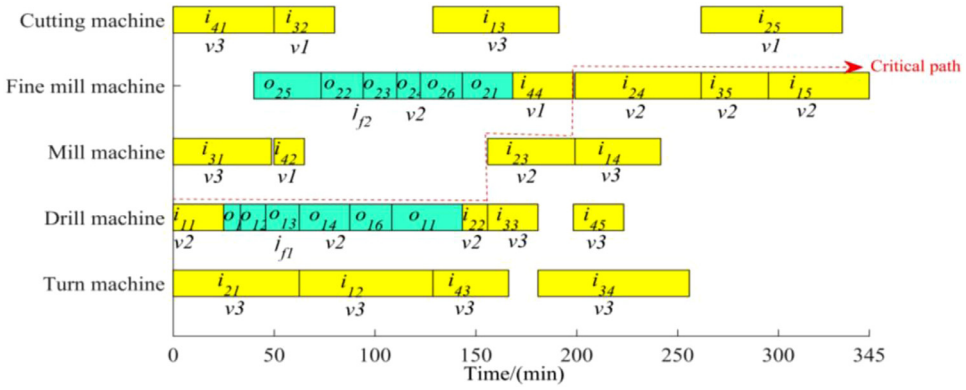


**Fig. 3.** Gantt chart of Pareto frontier solution A*

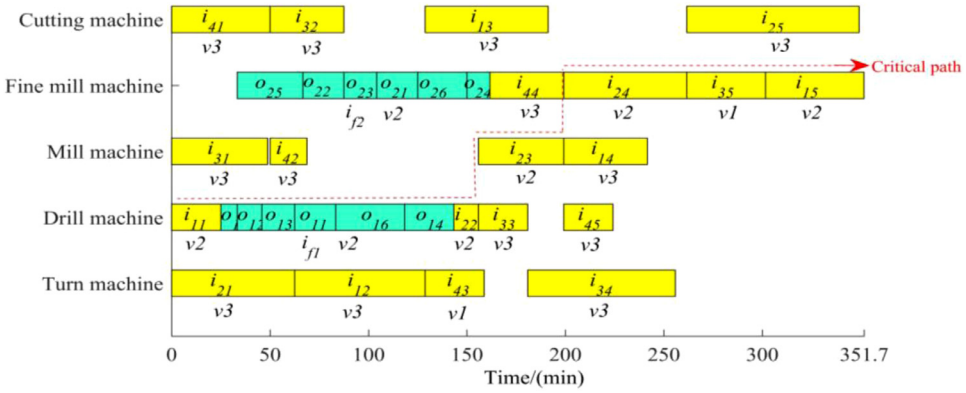


**Fig. 4.** Gantt chart of Pareto frontier solution C*

two columns in Table 2. Based on the reported results, the signal-to-noise ratio (SNR) main effect plots are drawn and shown in Fig. 5. Since bigger SNR indicates the better robustness of the parameter combination, the best parameter combination is {population size=40, mixing ratio=0.8, weight dimension=30, state size=20, learning rate=0.001, discount factor=0.8, greedy factor=0.85}. In addition, the experimental results reveal that a better performance is achieved when each weight in the weigh vectors takes a random value from the interval [0.2, 0.8].

To make a fair comparison, a similar parameter calibration procedure is adopted for all comparison MOEAs including NSGA-II [15], MOEA/D [17], multi-objective artificial bee colony algorithm (MOABC) [41], multi-objective population extremal optimization (MOEO) [42], multi-objective imperialist competitive algorithm with variables neighborhood search (MOICA) [43] and multi-objective genetic algorithm with manifold learning-inspired mating strategy (MOGA) [44]. Table 3 reports the final recommended parameter values.

### 5.4. Effectiveness of Bi-criteria selection

To evaluate the performance of Bi-criteria selection, three variants of NSGA-II with different selection mechanisms are designed and compared. For simplicity, they are named NSGA-II-RS, NSGA-II-AS and NSGA-II-BS. The selection mechanism used in NSGA-II is Pareto-based criterion; the adopted selection criteria in the three variants are R2 indicator-based, aggregation indicator-based and Bi-criteria-based, respectively. For the Bi-criteria-based selection, the earlier stage selects the excellent solutions into the next iteration according to the Pareto-based criterion, while the later stage adopts the aggregation indicator-based criterion.

For a fair comparison, each algorithm is executed ten times under the same stopping criterion. Table 4 reports the numerical results. From Table 4, the *HVR* and *IGD* values obtained by NSGA-II-RS are far worse than those by the others. This indicates that the R2 indicator-based criterion is not suitable for a multi-objective comprehensive evaluation and hence is ignored in the following analysis. To show the comparison results of the rest three algorithms more intuitively, Fig. 6 reports the 95% confidence interval plots of *HVR, IGD* and *C* metric. Clearly, NSGA-II-BS shows a higher superiority over NSGA-II-AS and NSGA-II in both *HVR* and *IGD*; and in terms of *C* metric, NSGA-II-BS Pareto dominates NSGA-II-AS and NSGA-II in more cases. All these demonstrate that the proposed Bi-criteria-based selection improves the performances of NSGA-II. The reason is that Bi-criteria-based selection combines the good diversity provided by the Pareto-based selection and the faster convergence given by aggregation indicator-based selection.

### 5.5. Effectiveness of Q-learning-based hyper-heuristic

In each iteration, the Q-learning-based hyper-heuristic recommends the best optimizer from {grey wolf operator, Jaya operator, crossover operator} to reproduce new solutions. To evaluate the performance of Q-learning-based hyper-heuristic, QHH-BS is compared with the other four MOEAs with different reproduction methods. These MOEAs are named JAYA-BS, CO-BS, GWO-BS and SHH-BS. Note that, Bi-criteria selection is utilized in all of these four comparisons. JAYA-BS, CO-BS, GWO-BS are single optimizer-based MOEAs and reproduce new solutions by the grey wolf operator, Jaya operator and crossover operator, respectively. SHH-BS is a simple hyper-heuristic, in which the three above optimizers are randomly selected and applied to achieve reproduction in each iteration. These four MOEAs and QHH-BS adopt the same encoding scheme and stopping criterion. Table 5 details the numerical results of *HVR, IGD* and *C* metric. Fig. 7 reports the corresponding confidence interval plots.

From Table 5 and Fig. 7, QHH-BS achieves the largest *HVR* in all cases. Regarding the *IGD* criterion, QHH-BS outperforms the three-single optimizer-based MOEAs in all cases and performs better than SHH-BS in 34/54 cases. If sequencing the four comparisons and QHH-BS in terms of *IGD* and *HVR*, QHH-BS ranks first, followed by SHH-BS and then the single optimizer-based MOEAs. As for the *C* metric, most of the Pareto frontier solutions achieved by the four comparisons are dominated by

**Table 2**
Orthogonal table of QHH-BS.

| No. | Factors | | | | | | | Response | |
|---|---|---|---|---|---|---|---|---|---|
| | Population size | Mixing ratio | Weight dimension | State size | Learning rate | Discount factor | Greedy factor | Avg *IGD* | Avg *HVR* |
| 1 | 30 | 0.7 | 20 | 10 | 0.0001 | 0.7 | 0.80 | 0.3627 | 0.2760 |
| 2 | 30 | 0.7 | 30 | 10 | 0.0001 | 0.8 | 0.85 | 0.0608 | 0.5817 |
| 3 | 30 | 0.7 | 40 | 10 | 0.0001 | 0.9 | 0.90 | 0.0802 | 0.5701 |
| 4 | 30 | 0.8 | 20 | 20 | 0.001 | 0.7 | 0.80 | 0.0410 | 0.6790 |
| 5 | 30 | 0.8 | 30 | 20 | 0.001 | 0.8 | 0.85 | 0.0729 | 0.5211 |
| 6 | 30 | 0.8 | 40 | 20 | 0.001 | 0.9 | 0.90 | 0.0709 | 0.4999 |
| 7 | 30 | 0.9 | 20 | 30 | 0.01 | 0.7 | 0.80 | 0.1575 | 0.3694 |
| 8 | 30 | 0.9 | 30 | 30 | 0.01 | 0.8 | 0.85 | 0.0605 | 0.5886 |
| 9 | 30 | 0.9 | 40 | 30 | 0.01 | 0.9 | 0.90 | 0.1729 | 0.4130 |
| 10 | 40 | 0.7 | 40 | 20 | 0.01 | 0.7 | 0.85 | 0.0641 | 0.5629 |
| 11 | 40 | 0.7 | 20 | 20 | 0.01 | 0.8 | 0.90 | 0.0337 | 0.6264 |
| 12 | 40 | 0.7 | 30 | 20 | 0.01 | 0.9 | 0.80 | 0.0654 | 0.5280 |
| 13 | 40 | 0.8 | 40 | 30 | 0.0001 | 0.7 | 0.85 | 0.0777 | 0.5649 |
| 14 | 40 | 0.8 | 20 | 30 | 0.0001 | 0.8 | 0.90 | 0.0705 | 0.5643 |
| 15 | 40 | 0.8 | 30 | 30 | 0.0001 | 0.9 | 0.80 | 0.0479 | 0.6342 |
| 16 | 40 | 0.9 | 40 | 10 | 0.001 | 0.7 | 0.85 | 0.0858 | 0.4980 |
| 17 | 40 | 0.9 | 20 | 10 | 0.001 | 0.8 | 0.90 | 0.0615 | 0.5203 |
| 18 | 40 | 0.9 | 30 | 10 | 0.001 | 0.9 | 0.80 | 0.0702 | 0.5508 |
| 19 | 50 | 0.7 | 30 | 30 | 0.001 | 0.7 | 0.90 | 0.0680 | 0.5758 |
| 20 | 50 | 0.7 | 40 | 30 | 0.001 | 0.8 | 0.80 | 0.0624 | 0.6056 |
| 21 | 50 | 0.7 | 20 | 30 | 0.001 | 0.9 | 0.85 | 0.0631 | 0.5973 |
| 22 | 50 | 0.8 | 30 | 10 | 0.01 | 0.7 | 0.90 | 0.0729 | 0.5297 |
| 23 | 50 | 0.8 | 40 | 10 | 0.01 | 0.8 | 0.80 | 0.0533 | 0.5490 |
| 24 | 50 | 0.8 | 20 | 10 | 0.01 | 0.9 | 0.85 | 0.0744 | 0.5305 |
| 25 | 50 | 0.9 | 30 | 20 | 0.0001 | 0.7 | 0.90 | 0.0860 | 0.4812 |
| 26 | 50 | 0.9 | 40 | 20 | 0.0001 | 0.8 | 0.80 | 0.1093 | 0.4084 |
| 27 | 50 | 0.9 | 20 | 20 | 0.0001 | 0.9 | 0.85 | 0.0694 | 0.5824 |

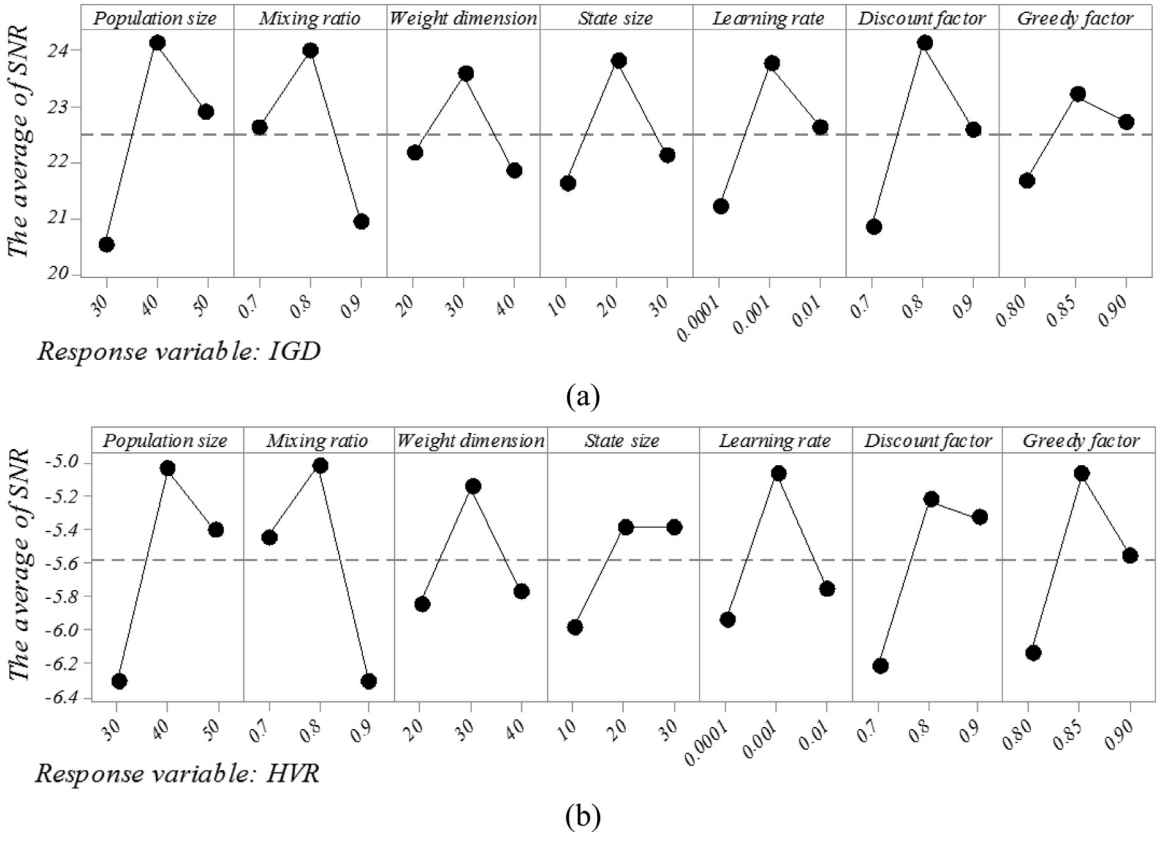


**Fig. 5.** Signal-to-noise ratio main effect plots with the response variables *IGD* (a) and *HVR* (b) respectively.

**Table 3**
Parameter settings of the comparison MOEAs.

| Parameters | Levels | NSGA-II | MOEA/D | MOABC | MOEO | MOICA | MOGA |
|---|---|---|---|---|---|---|---|
| Population size | 30, 40, 50 | 30 | 50 | 50 | 40 | 50 | 40 |
| Crossover rate | 0.85, 0.9, 0.95 | 0.9 | 0.95 | - | - | - | 0.95 |
| Mutation rate | 0.06, 0.08, 0.1 | 0.08 | 0.1 | - | - | - | 0.1 |
| Neighborhood size | 6, 8, 10 | - | 6 | - | - | - | - |
| Number of trials | 6, 8, 10 | - | - | 8 | - | | - |
| The shape factor of multi-non-uniform mutation | 0.4, 0.6, 0.8 | - | - | - | 0.6 | - | - |
| Number of imperialists | 6, 8, 10 | - | - | - | - | 8 | - |
| Revolution probability | 0.3, 0.35, 0.4 | - | - | - | - | 0.35 | - |

**Table 4**
Comparison results of NSGA-II algorithms with different selection criteria.

| Case | NSGA-II-RS *HVR/IGD/C*(MA-B, MA-R)/*C*(MA-R, MA-B) | NSGA-II-AS *HVR/IGD/C*(MA-B, MA-A)/*C*(MA-BJ, MA-B) | NSGA-II *HVR/IGD/C*(MA-B,MA-P)/ *C*(MA-P, MA-B) | NSGA-II-BS *HVR/IGD* |
|---|---|---|---|---|
| La21-F1 | 0.42/0.162/1/0 | 0.699/0.117/0.437/0.207 | 0.85/0.053/0.514/0.534 | **0.903/0.033** |
| La22-F1 | 0.445/0.125/0.933/0 | 0.817/0.053/0.463/0.555 | **0.863**/0.05/0.38/0.516 | 0.801/**0.047** |
| La23-F1 | 0.405/0.173/1/0 | 0.767/0.117/0.563/0.274 | 0.703/0.094/0.663/0.291 | **0.933/0.045** |
| La24-F1 | 0.443/0.118/0.986/0 | 0.83/0.051/0.206/0.228 | 0.849/0.055/0.472/0.314 | **0.888/0.038** |
| La25-F1 | 0.477/0.117/0.944/0.114 | 0.835/0.041/0.347/0.449 | **0.841**/0.052/0.279/0.696 | 0.796/**0.031** |
| La26-F1 | 0.424/0.073/0.971/0 | 0.806/0.039/0.564/0.181 | 0.721/0.036/0.581/0.19 | **0.844/0.027** |
| La27-F1 | 0.546/0.096/0.868/0.125 | **0.897**/0.06/0.18/0.495 | 0.787/**0.045**/0.457/0.403 | 0.764/0.046 |
| La28-F1 | 0.381/0.122/1/0 | 0.659/0.079/0.61/0.233 | 0.838/0.069/0.412/0.348 | **0.847/0.033** |
| La29-F1 | 0.413/0.155/0.906/0 | 0.735/0.084/0.549/0.209 | **0.908/0.044**/0.309/0.399 | 0.758/0.073 |
| La21-F2 | 0.447/0.083/0.884/0.01 | 0.721/0.063/0.383/0.426 | **0.845/0.024**/0.207/0.522 | 0.757/0.044 |
| La22-F2 | 0.447/0.103/0.86/0.096 | 0.804/0.081/0.167/0.354 | **0.808/0.06**/0.388/0.546 | 0.755/0.064 |
| La23-F2 | 0.395/0.173/0.889/0.01 | 0.796/0.071/0.117/0.632 | **0.88/0.058**/0.339/0.67 | 0.695/0.093 |
| La24-F2 | 0.355/0.144/1/0 | 0.769/**0.052**/0.375/0.18 | 0.783/0.059/0.513/0.303 | **0.9**/0.053 |
| La25-F2 | 0.408/0.133/0.91/0.089 | 0.665/0.081/0.933/0 | 0.815/0.07/0.362/0.214 | **0.886/0.039** |
| La26-F2 | 0.48/0.11/0.9/0.029 | 0.813/0.091/0.523/0.199 | 0.761/**0.044**/0.567/0.043 | **0.851**/0.055 |
| La27-F2 | 0.331/0.212/1/0 | 0.672/0.105/0.435/0.436 | 0.805/**0.055**/0.525/0.277 | **0.825**/0.077 |
| La28-F2 | 0.422/0.163/0.843/0.032 | **0.879**/0.075/0.15/0.743 | 0.805/**0.06**/0.235/0.525 | 0.689/0.096 |
| La29-F2 | 0.456/0.112/0.828/0.071 | **0.907**/0.051/0.131/0.692 | 0.815/**0.04**/0.313/0.45 | 0.795/0.068 |
| La21-F3 | 0.434/0.179/0.962/0 | 0.796/0.093/0.399/0.299 | 0.762/0.093/0.61/0.156 | **0.851/0.03** |
| La22-F3 | 0.407/0.078/0.973/0.014 | **0.806**/0.043/0.128/0.43 | 0.754/**0.035**/0.277/0.268 | 0.738/0.036 |
| La23-F3 | 0.431/0.161/0.98/0 | **0.909/0.034**/0.3/0.283 | 0.819/0.048/0.611/0.205 | 0.875/0.034 |
| La24-F3 | 0.471/0.109/0.903/0.032 | 0.848/0.062/0.301/0.39 | **0.87/0.031**/0.225/0.566 | 0.811/0.053 |
| La25-F3 | 0.42/0.109/0.988/0 | **0.901**/0.048/0.348/0.284 | 0.714/0.058/0.596/0.051 | 0.887/**0.04** |
| La26-F3 | 0.47/0.119/0.987/0 | **0.884/0.034**/0.239/0.374 | 0.86/0.045/0.414/0.264 | 0.802/0.041 |
| La27-F3 | 0.396/0.14/0.933/0 | **0.911/0.019**/0.107/0.667 | 0.823/0.059/0.217/0.544 | 0.767/0.062 |
| La28-F3 | 0.415/0.172/0.962/0 | 0.803/0.092/0.51/0.314 | 0.724/0.103/0.763/0.276 | **0.862/0.05** |
| La29-F3 | 0.401/0.197/0.959/0 | **0.858/0.035**/0.217/0.591 | 0.785/0.066/0.209/0.248 | 0.816/0.062 |
| La21-F4 | 0.424/0.109/0.956/0 | **0.868**/0.04/0.339/0.293 | 0.828/0.047/0.274/0.388 | 0.807/**0.037** |
| La22-F4 | 0.494/0.1/0.915/0 | **0.856**/0.06/0.217/0.45 | 0.8/0.048/0.53/0.331 | 0.764/**0.04** |
| La23-F4 | 0.458/0.147/0.927/0 | 0.811/0.063/0.649/0.078 | 0.839/**0.03**/0.321/0.333 | **0.848**/0.046 |
| La24-F4 | 0.447/0.131/0.699/0.12 | 0.87/**0.045**/0.25/0.481 | **0.872/0.045**/0.184/0.695 | 0.762/0.088 |
| La25-F4 | 0.39/0.171/0.953/0 | 0.801/0.054/0.575/0.309 | 0.885/**0.051**/0.459/0.337 | **0.9**/0.07 |
| La26-F4 | 0.461/0.105/0.968/0 | 0.858/0.056/0.192/0.417 | **0.861**/0.065/0.439/0.437 | 0.834/**0.032** |
| La27-F4 | 0.499/0.2/0.931/0.05 | **0.915/0.023**/0.096/0.779 | 0.817/0.092/0.41/0.257 | 0.785/0.096 |
| La28-F4 | 0.461/0.121/0.842/0.055 | 0.798/0.085/0.339/0.099 | 0.799/**0.057**/0.384/0.371 | **0.805**/0.068 |
| La29-F4 | 0.488/0.147/0.908/0 | **0.864/0.042**/0.402/0.405 | 0.676/0.097/0.822/0.103 | **0.864**/0.072 |
| La21-F5 | 0.518/0.122/0.956/0.063 | 0.802/0.069/0.311/0.649 | **0.82/0.041**/0.357/0.55 | 0.766/0.054 |
| La22-F5 | 0.504/0.145/0.938/0.042 | 0.675/0.095/0.659/0.249 | 0.704/0.094/0.716/0.256 | **0.896/0.047** |
| La23-F5 | 0.483/0.148/0.77/0.078 | 0.768/0.108/0.328/0.262 | **0.898/0.049**/0.134/0.643 | 0.77/0.076 |
| La24-F5 | 0.521/0.128/0.883/0 | 0.85/0.08/0.562/0.06 | 0.79/0.069/0.481/0.242 | **0.853/0.029** |
| La25-F5 | 0.491/0.107/0.964/0 | 0.674/0.084/0.807/0.129 | 0.785/0.072/0.65/0.144 | **0.92/0.02** |
| La26-F5 | 0.498/0.108/0.861/0.094 | 0.666/0.064/0.65/0.174 | 0.694/0.08/0.711/0.333 | **0.872/0.039** |
| La27-F5 | 0.432/0.118/0.93/0 | 0.68/0.086/0.396/0.154 | 0.753/0.104/0.643/0.2 | **0.864/0.05** |
| La28-F5 | 0.465/0.132/0.914/0.053 | **0.844**/0.067/0.4/0.36 | 0.765/0.062/0.738/0.213 | 0.838/**0.042** |
| La29-F5 | 0.552/0.077/0.829/0 | **0.85**/0.08/0.307/0.387 | 0.765/0.061/0.523/0.248 | 0.814/**0.037** |
| La21-F6 | 0.456/0.109/0.933/0.052 | 0.81/0.049/0.309/0.204 | 0.802/0.049/0.546/0.084 | **0.924/0.028** |
| La22-F6 | 0.558/0.085/0.907/0 | 0.851/0.041/0.427/0.325 | 0.782/0.075/0.383/0.242 | **0.863/0.019** |
| La23-F6 | 0.561/0.113/0.87/0.038 | 0.746/**0.084**/0.631/0.246 | **0.807**/0.089/0.498/0.451 | **0.807**/0.087 |
| La24-F6 | 0.478/0.157/0.888/0.083 | 0.766/0.08/0.459/0.4 | **0.823**/0.059/0.428/0.362 | 0.812/**0.05** |
| La25-F6 | 0.529/0.115/0.828/0.037 | 0.724/0.069/0.565/0.258 | 0.816/**0.065**/0.544/0.263 | **0.817**/0.091 |
| La26-F6 | 0.465/0.132/0.662/0.16 | 0.768/0.087/0.23/0.302 | **0.826/0.039**/0.075/0.612 | 0.696/0.07 |
| La27-F6 | 0.577/0.079/0.922/0.082 | 0.73/0.072/0.375/0.348 | **0.868/0.032**/0.37/0.295 | 0.833/0.054 |
| La28-F6 | 0.575/0.127/0.834/0.1 | 0.782/0.074/0.355/0.358 | **0.85/0.061**/0.45/0.429 | 0.793/0.086 |
| La29-F6 | 0.553/0.147/0.87/0 | 0.764/0.056/0.524/0.394 | 0.785/0.076/0.722/0.122 | **0.825/0.047** |

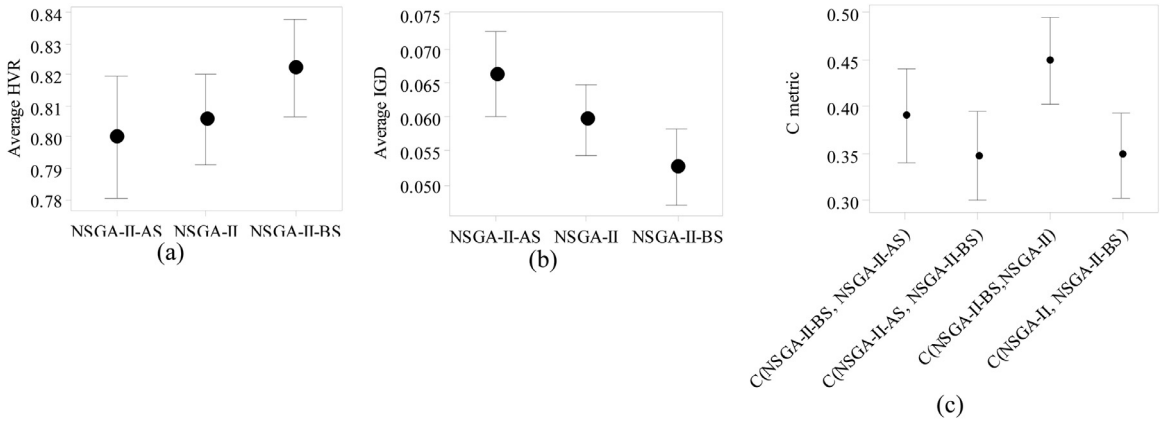


**Fig. 6.** 95% confidence interval plots for *HVR* (a), *IGD* (b) and *C* metric (c) of MOEAs with different selection criteria.

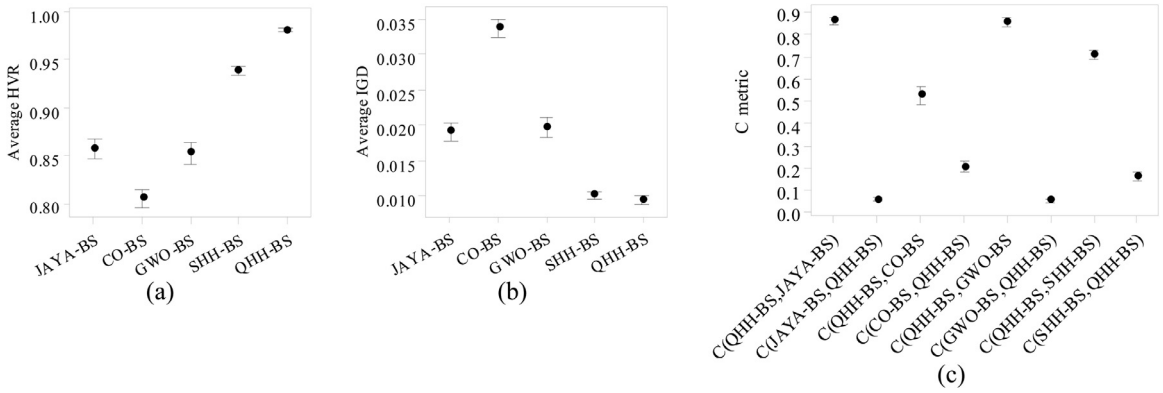


**Fig. 7.** 95% confidence interval plots for *HVR* (a), *IGD* (b) and *C* metric (c) of MOEAs with different optimizers.

solutions from QHH-BS. In contrast, only a very small proportion of solutions in QHH-BS are dominated by the four comparisons. Besides, individual optimizers perform differently in different cases. For example, JAYA-BS performs better than GWO-BS in case La21-F1, but it loses its superiority in case La24-F1.

To summarize, three findings are obtained. (1) Individual optimizers are sensitive to cases, and hence they are not robust enough to effectively deal with a wide range of cases. (2) Hyper-heuristics are more efficient and robust than single optimizer-based MOEAs. The reason is that each low-level optimizer behaves differently in exploration and exploitation, and hyper-heuristics benefit from the combination advantages of multiple optimizers and diversifies the searches in more promising regions of the solution space. (3) Q-learning-based hyper-heuristic shows significant superiority, because it adaptively selects an appropriate optimizer at each step in the search process so that the combination of multiple optimizers is more effective.

### 5.6. Validity of the proposed MILP model

In this section, three small-scaled cases are solved by GAMS/Cplex and QHH-BS to verify the validity of the proposed MILP model and QHH-BS. The processing time and processing machine of these cases are intercepted from the first three, four and five jobs of La01 and F1. The processing powers of the five involved machines are {20, 15, 20, 10, 18}. The processing energy consumption exponent α is 1.2. In addition, the speed level of each machine is set to be {1, 1.2, 0.8}.

Since the considered problem contains two objectives, the Epsilon-constraint method is applied in the solution process by GAMS/Cplex to obtain the Pareto front solutions. This method first obtains the lower bound of *TEC* ($TEC^l$) and the upper bound of *TEC* ($TEC^u$) by solving the MILP model with *TEC* objective and *Cmax* objective, respectively. Then, the interval $[TEC^l, TEC^u]$ is discretized into *K* equally subintervals. Finally, provided that the *TEC* value is less than the upper limit of each subinterval, the MILP model is solved with a single objective of minimizing *Cmax*. For simplicity, *K* takes a value of 50, and the MILP model is solved under 50 *TEC* values to obtain a series of *Cmax* values. Finally, the Pareto front solutions are achieved by removing the dominated solutions.

Through GAMS/Cplex, the *TEC* intervals of the three cases are [21278, 23076], [31972, 34673] and [41465, 44982]. The accumulated CPU times for 50 runs of the MILP model are 85 s, 364 s and 4360 s respectively. Fig. 8 shows the Pareto frontier plots obtained by the MILP model and QHH-BS. Note that, the CPU times by QHH-BS to obtain Pareto frontiers of the three cases are 100s, 150s and 200s, respectively.

Fig. 8 reports that both the MILP model and QHH-BS obtain a Pareto front set with well-distribution and well-convergence. Most Pareto front solutions obtained by the MILP model are the same as those by QHH-BS, while the rest of them are extremely close. These indicate that both the MILP model and QHH-BS are effective in finding Pareto front solutions of small-scaled problems. However, as the size of problems increases, the running time of GAMS/Cplex grows exponentially. Hence, for large-scaled cases, QHH-BS is more suitable to obtain Pareto frontier solutions in a shorter time.

### 5.7. Comparison of QHH-BS with state-of-the-art MOEAs

To further demonstrate the superiority of QHH-BS, comparison experiments are performed with six competitive MOEAs including NSGA-II [15], MOEA/D [17], MOABC [41], MOEO [42], MOICA [43] and MOGA [44]. The reason for choosing these MOEAs is two-fold. (1) NSGA-II and MOEA/D are commonly utilized as baseline algorithms to evaluate

**Table 5**
Comparison results of MOEAs with different optimizers.

| Case | JAYA-BS<br>*HVR/IGD/C(QHH-BS,JAYA-BS)/C(JAYA-BS,QHH-BS)* | CO-BS<br>*HVR/IGD/C(QHH-BS, CO-BS) /C(CO-BS,QHH-BS)* | GWO-BS<br>*HVR/IGD/C(QHH-BS,GWO-BS)/C(GWO-BS,QHH-BS)* | SHH-BS<br>*HVR/IGD/C(QHH-BS,SHH-BS)/C(SHH-BS,QHH-BS)* | QHH-BS<br>*HVR/IGD* |
|---|---|---|---|---|---|
| La21-F1 | 0.846/0.017/0.884/0.028 | 0.809/0.027/0.355/0.316 | 0.838/0.02/0.916/0.024 | 0.911/0.011/0.836/0.06 | **0.988/0.008** |
| La22-F1 | 0.897/0.014/0.864/0.047 | 0.819/0.034/0.343/0.26 | 0.876/0.015/0.958/0.008 | 0.922/0.01/0.813/0.062 | **0.99/0.009** |
| La23-F1 | 0.888/0.016/0.834/0.101 | 0.799/0.032/0.386/0.191 | 0.838/0.021/0.907/0.029 | 0.931/0.01/0.745/0.155 | **0.985/0.01** |
| La24-F1 | 0.828/0.022/0.858/0.042 | 0.837/0.037/0.609/0.401 | 0.865/0.019/0.902/0.03 | 0.963/**0.007**/0.5/0.285 | **0.977**/0.01 |
| La25-F1 | 0.899/0.015/0.775/0.039 | 0.801/0.043/0.601/0.112 | 0.828/0.024/0.833/0.006 | 0.935/**0.011**/0.775/0.063 | **0.973**/0.017 |
| La26-F1 | 0.87/0.017/0.932/0.023 | 0.782/0.035/0.464/0.186 | 0.84/0.02/0.923/0.029 | 0.936/0.008/0.794/0.056 | **0.994/0.008** |
| La27-F1 | 0.91/0.013/0.735/0.158 | 0.774/0.034/0.667/0.098 | 0.824/0.02/0.914/0.008 | 0.938/**0.008**/0.656/0.168 | **0.971**/0.011 |
| La28-F1 | 0.879/0.016/0.897/0.061 | 0.785/0.038/0.303/0.259 | 0.895/0.015/0.72/0.137 | 0.939/0.008/0.676/0.201 | **0.975/0.007** |
| La29-F1 | 0.871/0.018/0.769/0.11 | 0.805/0.034/0.67/0.338 | 0.865/0.017/0.716/0.158 | 0.937/**0.008**/0.697/0.19 | **0.971**/0.009 |
| La21-F2 | 0.865/0.016/0.878/0.02 | 0.838/0.023/0.686/0.391 | 0.857/0.017/0.881/0.025 | 0.938/**0.008**/0.736/0.099 | **0.982**/0.009 |
| La22-F2 | 0.861/0.018/0.861/0.044 | 0.816/0.031/0.34/0.33 | 0.841/0.019/0.782/0.056 | 0.937/0.01/0.702/0.156 | **0.974/0.008** |
| La23-F2 | 0.878/0.016/0.771/0.118 | 0.79/0.04/0.393/0.237 | 0.883/0.016/0.764/0.08 | 0.949/0.01/0.59/0.237 | **0.979/0.008** |
| La24-F2 | 0.872/0.018/0.833/0.069 | 0.8/0.036/0.332/0.207 | 0.868/0.019/0.82/0.111 | 0.943/0.01/0.713/0.159 | **0.974/0.009** |
| La25-F2 | 0.876/0.018/0.88/0.059 | 0.807/0.035/0.465/0.197 | 0.83/0.02/0.895/0.029 | 0.944/0.01/0.709/0.214 | **0.975/0.007** |
| La26-F2 | 0.892/0.013/0.886/0.029 | 0.756/0.035/0.435/0.055 | 0.852/0.017/0.854/0.085 | 0.922/0.011/0.801/0.083 | **0.991/0.008** |
| La27-F2 | 0.91/0.013/0.764/0.1 | 0.783/0.036/0.413/0.232 | 0.865/0.017/0.794/0.149 | 0.941/**0.008**/0.594/0.25 | **0.983**/0.012 |
| La28-F2 | 0.916/0.012/0.721/0.091 | 0.8/0.032/0.304/0.291 | 0.855/0.018/0.803/0.056 | 0.931/**0.01**/0.634/0.283 | **0.97**/0.012 |
| La29-F2 | 0.849/0.02/0.728/0.16 | 0.813/0.035/0.382/0.255 | 0.858/0.02/0.868/0.047 | 0.928/0.011/0.633/0.249 | **0.97/0.009** |
| La21-F3 | 0.751/0.03/0.921/0.019 | 0.808/0.026/0.552/0.22 | 0.765/0.029/0.938/0.028 | 0.921/0.012/0.756/0.123 | **0.984/0.006** |
| La22-F3 | 0.822/0.027/0.916/0.024 | 0.833/0.029/0.722/0.181 | 0.81/0.029/0.905/0.032 | 0.928/0.013/0.829/0.103 | **0.985/0.006** |
| La23-F3 | 0.823/0.022/0.95/0.011 | 0.804/0.031/0.667/0.142 | 0.807/0.024/0.925/0.008 | 0.921/0.013/0.884/0.022 | **0.992/0.008** |
| La24-F3 | 0.827/0.023/0.896/0.054 | 0.822/0.033/0.487/0.284 | 0.826/0.024/0.913/0.033 | 0.943/0.01/0.694/0.212 | **0.969/0.01** |
| La25-F3 | 0.768/0.03/0.928/0.046 | 0.793/0.032/0.647/0.201 | 0.754/0.034/0.884/0.052 | 0.901/0.014/0.686/0.14 | **0.977/0.007** |
| La26-F3 | 0.857/0.021/0.847/0.018 | 0.768/0.034/0.634/0.134 | 0.83/0.023/0.867/0.027 | 0.921/0.012/0.768/0.121 | **0.974/0.009** |
| La27-F3 | 0.82/0.028/0.883/0.046 | 0.764/0.04/0.661/0.137 | 0.785/0.032/0.923/0.02 | 0.949/**0.014**/0.578/0.301 | **0.984**/0.015 |
| La28-F3 | 0.807/0.027/0.961/0.009 | 0.767/0.037/0.792/0.079 | 0.833/0.023/0.924/0.025 | 0.914/**0.013**/0.639/0.274 | 0.973/0.015 |
| La29-F3 | 0.872/0.018/0.898/0.042 | 0.771/0.036/0.605/0.124 | 0.818/0.023/0.897/0.04 | 0.95/0.01/0.677/0.234 | **0.987/0.007** |
| La21-F4 | 0.821/0.023/0.82/0.052 | 0.785/0.039/0.487/0.186 | 0.828/0.025/0.864/0.05 | 0.944/**0.011**/0.566/0.287 | **0.969**/0.012 |
| La22-F4 | 0.805/0.026/0.955/0.01 | 0.775/0.041/0.58/0.162 | 0.813/0.027/0.935/0.018 | 0.911/0.014/0.83/0.064 | **0.989/0.005** |
| La23-F4 | 0.833/0.022/0.901/0.016 | 0.808/0.041/0.64/0.101 | 0.826/0.025/0.938/0.017 | 0.925/0.013/0.743/0.129 | **0.99/0.008** |
| La24-F4 | 0.826/0.024/0.945/0.015 | 0.807/0.031/0.741/0.104 | 0.823/0.024/0.874/0.02 | 0.963/**0.008**/0.581/0.179 | **0.981**/0.012 |
| La25-F4 | 0.819/0.024/0.917/0.008 | 0.768/0.037/0.663/0.156 | 0.809/0.026/0.915/0.012 | 0.944/0.012/0.7/0.165 | **0.992/0.007** |
| La26-F4 | 0.804/0.023/0.943/0.02 | 0.747/0.035/0.734/0.095 | 0.778/0.027/0.885/0.051 | 0.912/0.013/0.675/0.21 | **0.986/0.006** |
| La27-F4 | 0.814/0.023/0.901/0.022 | 0.742/0.034/0.834/0.045 | 0.799/0.025/0.903/0.021 | 0.94/0.009/0.738/0.148 | **0.994/0.007** |
| La28-F4 | 0.86/0.018/0.853/0.063 | 0.782/0.03/0.514/0.139 | 0.854/0.019/0.808/0.056 | 0.929/0.011/0.708/0.174 | **0.979/0.008** |
| La29-F4 | 0.761/0.031/0.968/0.011 | 0.728/0.041/0.888/0.034 | 0.812/0.027/0.927/0.031 | 0.896/0.016/0.891/0.032 | **0.997/0.011** |
| La21-F5 | 0.898/0.017/0.828/0.091 | 0.832/0.036/0.501/0.139 | 0.89/0.015/0.855/0.052 | 0.931/0.01/0.793/0.088 | **0.991/0.009** |
| La22-F5 | 0.85/0.017/0.884/0.042 | 0.845/0.032/0.338/0.215 | 0.904/0.014/0.868/0.059 | 0.952/0.008/0.777/0.096 | **0.989/0.009** |
| La23-F5 | 0.865/0.017/0.873/0.037 | 0.853/0.033/0.345/0.221 | 0.921/0.012/0.75/0.077 | 0.959/**0.007**/0.642/0.145 | **0.983**/0.013 |
| La24-F5 | 0.911/0.014/0.814/0.099 | 0.865/0.027/0.462/0.231 | 0.879/0.017/0.871/0.029 | 0.946/0.008/0.752/0.112 | **0.983/0.007** |
| La25-F5 | 0.848/0.02/0.896/0.039 | 0.83/0.029/0.481/0.273 | 0.881/0.017/0.841/0.03 | 0.94/0.012/0.697/0.141 | **0.976/0.012** |
| La26-F5 | 0.913/0.014/0.821/0.02 | 0.827/0.034/0.693/0.298 | 0.88/0.016/0.875/0.029 | 0.944/**0.009**/0.736/0.101 | **0.98**/0.011 |
| La27-F5 | 0.906/0.015/0.839/0.051 | 0.828/0.032/0.699/0.309 | 0.906/0.014/0.813/0.066 | 0.953/**0.008**/0.62/0.244 | **0.977**/0.01 |
| La28-F5 | 0.877/0.018/0.771/0.096 | 0.8/0.037/0.394/0.189 | 0.889/0.015/0.74/0.08 | 0.938/0.011/0.68/0.178 | **0.97/0.009** |
| La29-F5 | 0.828/0.02/0.796/0.106 | 0.8/0.034/0.47/0.278 | 0.847/0.02/0.867/0.058 | 0.953/**0.008**/0.63/0.248 | **0.962**/0.01 |
| La21-F6 | 0.885/0.015/0.863/0.031 | 0.861/0.025/0.396/0.219 | 0.894/0.016/0.887/0.051 | 0.949/**0.008**/0.751/0.128 | **0.987**/0.009 |
| La22-F6 | 0.867/0.016/0.924/0.022 | 0.805/0.045/0.638/0.092 | 0.892/0.014/0.82/0.043 | 0.945/0.01/0.675/0.132 | **0.987/0.01** |
| La23-F6 | 0.896/0.016/0.891/0.021 | 0.859/0.026/0.469/0.192 | 0.894/0.015/0.785/0.05 | 0.959/**0.01**/0.692/0.173 | **0.985**/0.013 |
| La24-F6 | 0.839/0.019/0.92/0.026 | 0.859/0.026/0.473/0.213 | 0.903/0.013/0.856/0.024 | 0.952/0.008/0.791/0.079 | **0.992/0.008** |
| La25-F6 | 0.896/0.017/0.808/0.066 | 0.82/0.039/0.399/0.194 | 0.917/0.013/0.778/0.058 | 0.966/0.008/0.658/0.192 | **0.981/0.008** |
| La26-F6 | 0.899/0.014/0.804/0.079 | 0.828/0.035/0.322/0.287 | 0.932/0.011/0.714/0.086 | 0.969/**0.007**/0.691/0.14 | **0.986**/0.009 |
| La27-F6 | 0.851/0.018/0.915/0.039 | 0.833/0.032/0.371/0.182 | 0.913/0.013/0.87/0.057 | 0.957/0.009/0.694/0.162 | **0.994/0.007** |
| La28-F6 | 0.887/0.014/0.797/0.043 | 0.861/0.023/0.445/0.221 | 0.9/0.013/0.815/0.039 | 0.958/**0.008**/0.751/0.075 | **0.984**/0.01 |
| La29-F6 | 0.878/0.022/0.875/0.066 | 0.795/0.039/0.519/0.163 | 0.906/0.016/0.767/0.083 | 0.963/**0.01**/0.634/0.21 | **0.983**/0.013 |

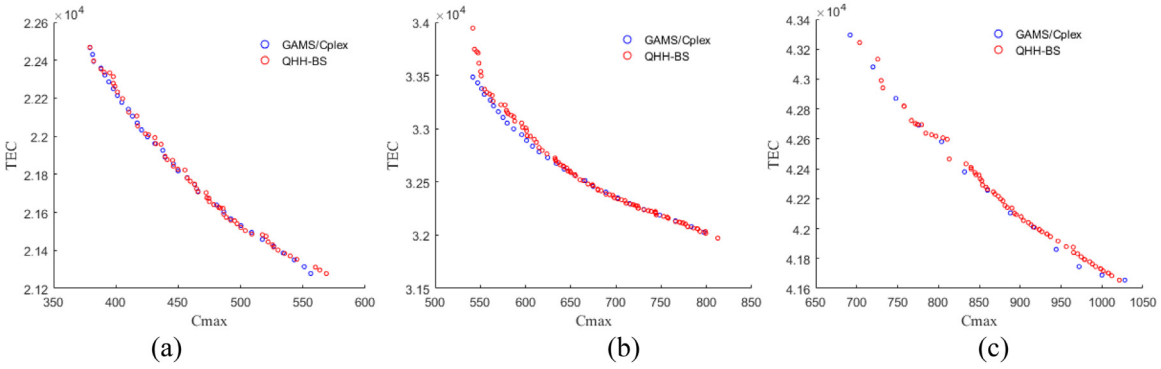


**Fig. 8.** Result Comparison between MILP model and QHH-BS in cases of (a) three jobs, (b) four jobs and (c) five jobs.

**Table 6**
Comparison results of QHH-BS and state-of-the-art MOEAs.

| Case | NSGA-II *HVR/IGD/C*(QHH-BS,NSGA-II)/*C*(NSHA-II,QHH-BS) | MOEA/D *HVR/IGD/C*(QHH-BS,MOEA/D)/ *C*(MOEA/D,QHH-BS) | MOABC *HVR/IGD/C*(QHH-BS,MOABC)/ *C*(MOABC,QHH-BS) | MOEO *HVR/IGD/C*(QHH-BS,MOEO)/ *C*(MOEO,QHH-BS) | MOICA *HVR/IGD/C*(QHH-BS,MOICA)/ *C*(MOICA,QHH-BS) | MOGA *HVR/IGD/C*(QHH-BS,MOGA)/ *C*(MOGA,QHH-BS) | QHH-BS *HVR/IGD* |
|---|---|---|---|---|---|---|---|
| La21-F1 | 0.82/0.03/0.55/0.24 | 0.57/0.04/1/0 | 0.58/0.04/1/0 | 0.67/0.03/1/0 | 0.8/0.03/0.72/0.09 | 0.6/0.04/1/0 | **0.99/0** |
| La22-F1 | 0.82/0.04/0.46/0.21 | 0.56/0.05/1/0 | 0.58/0.04/1/0 | 0.7/0.04/1/0 | 0.83/0.03/0.58/0.23 | 0.62/0.04/1/0 | **0.99/0.01** |
| La23-F1 | 0.81/0.04/0.51/0.19 | 0.55/0.05/1/0 | 0.58/0.04/1/0 | 0.68/0.04/0.98/0.01 | 0.81/0.04/0.67/0.1 | 0.62/0.04/1/0 | **0.99/0** |
| La24-F1 | 0.8/0.04/0.58/0.2 | 0.55/0.05/1/0 | 0.58/0.04/1/0 | 0.68/0.04/1/0 | 0.81/0.03/0.76/0.07 | 0.6/0.04/1/0 | **0.99/0.01** |
| La25-F1 | 0.82/0.04/0.49/0.19 | 0.58/0.04/1/0 | 0.6/0.04/1/0 | 0.68/0.04/1/0 | 0.82/0.03/0.79/0.1 | 0.6/0.04/1/0 | **0.99/0** |
| La26-F1 | 0.78/0.04/0.38/0.25 | 0.49/0.05/1/0 | 0.56/0.04/1/0 | 0.67/0.04/1/0 | 0.75/0.04/0.62/0.09 | 0.57/0.05/1/0 | **0.99/0.01** |
| La27-F1 | 0.81/0.04/0.48/0.2 | 0.53/0.04/1/0 | 0.56/0.04/1/0 | 0.64/0.04/1/0 | 0.75/0.03/0.94/0 | 0.6/0.04/1/0 | **0.99/0** |
| La28-F1 | 0.84/0.04/0.35/0.29 | 0.57/0.04/1/0 | 0.58/0.04/1/0 | 0.67/0.04/0.97/0.01 | 0.8/0.04/0.53/0.25 | 0.59/0.04/1/0 | **0.98/0.01** |
| La28-F1 | 0.81/0.04/0.7/0.17 | 0.55/0.05/1/0 | 0.57/0.04/1/0 | 0.64/0.04/1/0 | 0.78/0.03/0.6/0.2 | 0.59/0.04/1/0 | **0.99/0** |
| La21-F2 | 0.83/0.03/0.38/0.3 | 0.57/0.04/1/0 | 0.62/0.04/1/0 | 0.67/0.04/1/0 | 0.81/0.03/0.69/0.17 | 0.63/0.04/1/0 | **0.99/0.01** |
| La22-F2 | 0.81/0.04/0.38/0.19 | 0.54/0.05/1/0 | 0.57/0.05/1/0 | 0.66/0.04/1/0 | 0.79/0.04/0.74/0.1 | 0.59/0.05/1/0 | **1/0** |
| La23-F2 | 0.83/0.04/0.48/0.26 | 0.57/0.05/1/0 | 0.6/0.04/1/0 | 0.67/0.04/1/0 | 0.82/0.04/0.47/0.19 | 0.59/0.04/1/0 | **0.99/0.01** |
| La24-F2 | 0.78/0.05/0.83/0.03 | 0.57/0.05/1/0 | 0.59/0.05/1/0 | 0.7/0.04/1/0 | 0.82/0.04/0.61/0.09 | 0.59/0.05/1/0 | **0.99/0.01** |
| La25-F2 | 0.77/0.04/0.89/0.05 | 0.56/0.05/1/0 | 0.58/0.04/1/0 | 0.69/0.03/0.98/0.01 | 0.83/0.03/0.66/0.17 | 0.6/0.04/1/0 | **0.99/0** |
| La26-F2 | 0.83/0.04/0.54/0.19 | 0.52/0.05/1/0 | 0.59/0.04/1/0 | 0.68/0.03/1/0 | 0.77/0.03/0.66/0.13 | 0.6/0.04/1/0 | **0.99/0.01** |
| La27-F2 | 0.79/0.04/0.66/0.13 | 0.56/0.04/1/0 | 0.6/0.04/1/0 | 0.67/0.04/1/0 | 0.79/0.04/0.55/0.16 | 0.59/0.04/1/0 | **0.99/0.01** |
| La28-F2 | 0.79/0.04/0.59/0.19 | 0.55/0.04/1/0 | 0.57/0.04/1/0 | 0.68/0.03/1/0 | 0.78/0.03/0.65/0.11 | 0.61/0.04/1/0 | **1/0** |
| La29-F2 | 0.81/0.04/0.76/0.16 | 0.57/0.04/1/0 | 0.57/0.04/1/0 | 0.64/0.04/1/0 | 0.77/0.03/0.79/0.11 | 0.6/0.04/1/0 | **0.99/0** |
| La21-F3 | 0.79/0.03/0.79/0.1 | 0.56/0.04/1/0 | 0.57/0.04/1/0 | 0.69/0.03/1/0 | 0.8/0.02/0.97/0.01 | 0.58/0.04/1/0 | **0.99/0.01** |
| La22-F3 | 0.8/0.04/0.8/0.14 | 0.53/0.05/1/0 | 0.56/0.05/1/0 | 0.66/0.04/1/0 | 0.81/0.03/0.79/0.1 | 0.57/0.05/1/0 | **0.99/0.01** |
| La23-F3 | 0.81/0.03/0.75/0.14 | 0.54/0.05/1/0 | 0.57/0.04/1/0 | 0.67/0.03/0.99/0.01 | 0.79/0.03/0.89/0.04 | 0.6/0.04/1/0 | **0.99/0.01** |
| La24-F3 | 0.8/0.04/0.7/0.06 | 0.54/0.05/1/0 | 0.56/0.04/1/0 | 0.64/0.04/1/0 | 0.79/0.03/0.74/0.09 | 0.58/0.04/1/0 | **0.99/0.01** |
| La25-F3 | 0.79/0.03/0.78/0.1 | 0.51/0.05/1/0 | 0.54/0.05/1/0 | 0.61/0.04/1/0 | 0.81/0.03/0.66/0.17 | 0.56/0.04/1/0 | **0.99/0** |
| La26-F3 | 0.77/0.03/0.81/0.08 | 0.5/0.06/0.97/0 | 0.52/0.05/1/0 | 0.63/0.04/1/0 | 0.77/0.03/0.81/0.04 | 0.53/0.05/0.99/0 | **0.99/0.01** |
| La27-F3 | 0.77/0.03/0.68/0.08 | 0.52/0.05/1/0 | 0.54/0.04/1/0 | 0.69/0.03/0.99/0 | 0.8/0.03/0.63/0.15 | 0.57/0.04/1/0 | **0.99/0.01** |
| La28-F3 | 0.8/0.04/0.52/0.25 | 0.55/0.05/1/0 | 0.56/0.05/1/0 | 0.64/0.04/1/0 | 0.79/0.03/0.73/0.12 | 0.59/0.04/0.99/0 | **0.99/0.01** |
| La29-F3 | 0.78/0.04/0.61/0.15 | 0.51/0.06/1/0 | 0.55/0.05/1/0 | 0.65/0.04/1/0 | 0.8/0.04/0.59/0.19 | 0.57/0.05/1/0 | **0.98/0.01** |
| La21-F4 | 0.79/0.03/0.82/0.07 | 0.55/0.05/1/0 | 0.53/0.05/1/0 | 0.65/0.04/0.99/0 | 0.8/0.03/0.9/0.04 | 0.57/0.04/0.96/0 | **0.99/0.01** |
| La22-F4 | 0.79/0.03/0.83/0.09 | 0.53/0.05/1/0 | 0.56/0.05/1/0 | 0.65/0.04/0.99/0 | 0.77/0.03/0.84/0.06 | 0.59/0.05/1/0 | **1/0** |
| La23-F4 | 0.79/0.04/0.7/0.1 | 0.53/0.06/1/0 | 0.56/0.05/1/0 | 0.68/0.04/1/0 | 0.81/0.03/0.66/0.13 | 0.58/0.05/1/0 | **0.99/0.01** |
| La24-F4 | 0.8/0.04/0.72/0.15 | 0.55/0.06/1/0 | 0.57/0.05/1/0 | 0.65/0.05/0.91/0.03 | 0.83/0.04/0.59/0.24 | 0.57/0.05/1/0 | **0.99/0.01** |
| La25-F4 | 0.79/0.04/0.75/0.11 | 0.55/0.05/1/0 | 0.57/0.05/1/0 | 0.68/0.04/0.98/0.01 | 0.82/0.04/0.74/0.11 | 0.59/0.05/1/0 | **0.99/0.01** |
| La26-F4 | 0.76/0.04/0.89/0.09 | 0.49/0.06/1/0 | 0.55/0.05/1/0 | 0.67/0.04/1/0 | 0.77/0.03/0.94/0.04 | 0.57/0.04/1/0 | **1/0.01** |
| La27-F4 | 0.73/0.04/0.96/0.01 | 0.51/0.05/1/0 | 0.55/0.05/1/0 | 0.65/0.04/1/0 | 0.77/0.03/0.77/0.11 | 0.54/0.05/1/0 | **0.99/0.01** |
| La28-F4 | 0.79/0.04/0.88/0.07 | 0.54/0.05/1/0 | 0.57/0.05/1/0 | 0.65/0.04/1/0 | 0.79/0.04/0.73/0.16 | 0.59/0.05/1/0 | **0.99/0.01** |
| La29-F4 | 0.77/0.03/0.82/0.06 | 0.53/0.05/1/0 | 0.54/0.05/1/0 | 0.65/0.04/1/0 | 0.8/0.03/0.74/0.17 | 0.58/0.05/1/0 | **0.99/0** |
| La21-F5 | 0.82/0.04/0.65/0.11 | 0.6/0.05/1/0 | 0.67/0.03/1/0 | 0.73/0.03/1/0 | 0.85/0.02/0.76/0.12 | 0.64/0.04/1/0 | **1/0** |
| La22-F5 | 0.84/0.04/0.56/0.17 | 0.63/0.04/1/0.02 | 0.68/0.03/1/0 | 0.77/0.03/0.96/0.03 | 0.88/0.02/0.55/0.23 | 0.68/0.03/0.98/0 | **0.99/0.01** |
| La23-F5 | 0.83/0.03/0.57/0.19 | 0.59/0.05/1/0 | 0.67/0.03/1/0 | 0.77/0.02/1/0 | 0.85/0.02/0.66/0.16 | 0.66/0.04/1/0 | **0.99/0.01** |
| La24-F5 | 0.87/0.04/0.47/0.21 | 0.63/0.05/1/0 | 0.71/0.03/1/0 | 0.78/0.03/0.95/0.03 | 0.88/0.03/0.51/0.23 | 0.66/0.04/0.93/0.03 | **0.99/0** |
| La25-F5 | 0.83/0.03/0.63/0.11 | 0.57/0.05/1/0 | 0.66/0.04/1/0 | 0.74/0.03/0.99/0 | 0.88/0.02/0.57/0.18 | 0.67/0.04/1/0 | **0.99/0.01** |
| La26-F5 | 0.81/0.04/0.54/0.11 | 0.54/0.05/1/0 | 0.62/0.04/1/0 | 0.69/0.04/1/0 | 0.81/0.03/0.76/0.08 | 0.63/0.04/1/0 | **0.99/0.01** |
| La27-F5 | 0.8/0.03/0.6/0.15 | 0.55/0.05/1/0 | 0.66/0.03/1/0 | 0.71/0.03/0.98/0 | 0.85/0.02/0.68/0.15 | 0.61/0.04/1/0 | **0.99/0** |
| La28-F5 | 0.77/0.04/0.85/0.04 | 0.59/0.04/1/0 | 0.65/0.04/1/0 | 0.74/0.03/0.99/0.01 | 0.83/0.03/0.61/0.16 | 0.66/0.04/1/0 | **1/0** |
| La29-F5 | 0.82/0.03/0.87/0.03 | 0.59/0.05/1/0 | 0.67/0.04/1/0 | 0.73/0.04/1/0 | 0.87/0.02/0.59/0.21 | 0.64/0.04/1/0 | **0.99/0.01** |
| La21-F6 | 0.84/0.03/0.57/0.18 | 0.57/0.05/1/0 | 0.66/0.03/1/0 | 0.73/0.03/0.98/0 | 0.85/0.02/0.53/0.18 | 0.66/0.04/1/0 | **0.99/0** |
| La22-F6 | 0.83/0.04/0.51/0.09 | 0.64/0.04/1/0 | 0.72/0.03/1/0 | 0.75/0.03/0.98/0 | 0.86/0.03/0.64/0.12 | 0.68/0.04/1/0 | **0.99/0.01** |
| La23-F6 | 0.82/0.04/0.67/0.13 | 0.64/0.04/1/0 | 0.7/0.03/1/0 | 0.78/0.03/0.95/0.05 | 0.88/0.03/0.57/0.23 | 0.69/0.04/1/0 | **0.99/0** |
| La24-F6 | 0.83/0.04/0.46/0.21 | 0.59/0.05/1/0 | 0.69/0.03/1/0 | 0.78/0.03/0.92/0.05 | 0.87/0.02/0.58/0.22 | 0.63/0.04/1/0 | **0.99/0.01** |
| La25-F6 | 0.86/0.03/0.33/0.28 | 0.57/0.05/1/0 | 0.67/0.03/1/0 | 0.75/0.03/0.97/0 | 0.88/0.02/0.58/0.17 | 0.67/0.04/0.97/0 | **0.98/0.01** |
| La26-F6 | 0.83/0.04/0.43/0.14 | 0.61/0.04/1/0 | 0.7/0.03/1/0 | 0.75/0.03/0.96/0.02 | 0.86/0.03/0.61/0.15 | 0.67/0.04/1/0 | **0.99/0** |
| La27-F6 | 0.84/0.03/0.5/0.22 | 0.6/0.04/1/0 | 0.68/0.03/1/0 | 0.71/0.03/1/0 | 0.84/0.02/0.74/0.12 | 0.66/0.04/1/0 | **0.99/0** |
| La28-F6 | 0.85/0.03/0.4/0.21 | 0.56/0.04/1/0 | 0.68/0.03/1/0 | 0.77/0.03/0.9/0.02 | 0.86/0.03/0.39/0.19 | 0.65/0.04/0.98/0 | **0.99/0.01** |
| La29-F6 | 0.77/0.03/0.85/0.11 | 0.58/0.05/1/0 | 0.66/0.03/1/0 | 0.74/0.03/1/0 | 0.85/0.03/0.59/0.2 | 0.64/0.04/1/0 | **0.99/0** |

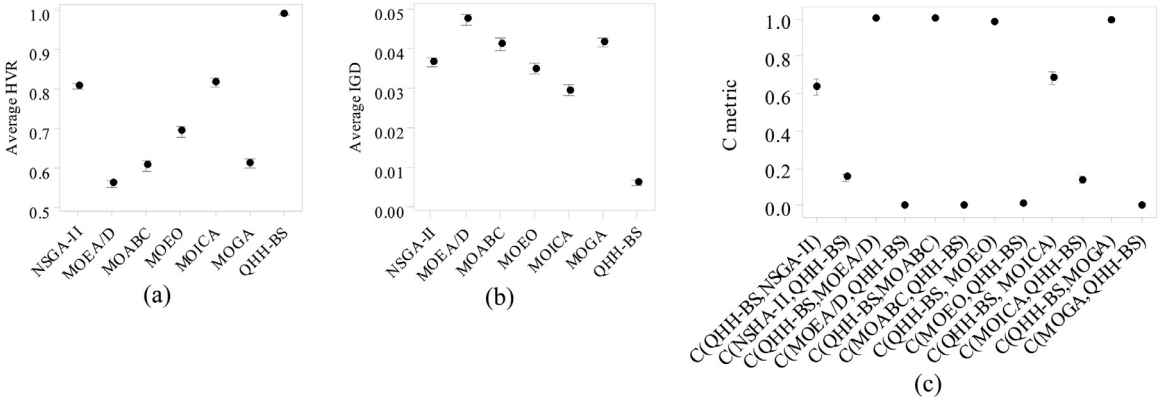


**Fig. 9.** 95% confidence interval plots for *HVR* (a), *IGD* (b) and *C* metric (c) of QHH-BS and state-of-the-art MOEAs.

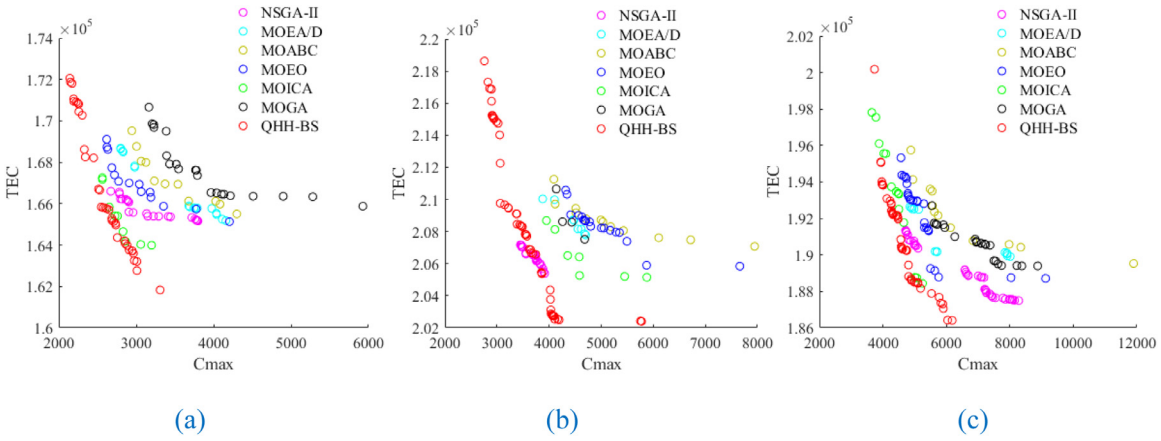


**Fig. 10.** Pareto frontier plots by QHH-BS and comparison MOEAs on (a) La21-F2, (b) La21-F4 and (c) La21-F6.

**Table 7**
Mean Rank and *P*-values derived by the Friedman test.

| Friedman test | NSGA-II | MOEA/D | MOABC | MOEO | MOICA | MOGA | QHH-BS | *p*-value |
|---|---|---|---|---|---|---|---|---|
| Rank of *IGD* | 3.98 | 6.96 | 5.28 | 3.37 | 2.15 | 5.26 | 1 | 4.4135E−59 |
| Rank of *HVR* | 5.35 | 1.02 | 2.31 | 4 | 5.65 | 2.67 | 7 | 1.3276E−64 |

the performance of the newly proposed MOEAs. (2) MOABC, MOICA, MOGA and MOEO have been proved to be effective in solving energy-aware shop scheduling problems or complex combinatorial optimization problems.

Table 6 and Fig. 9 report the comparison results in terms of *HVR, IGD* and *C* metric. It clearly shows that QHH-BS achieves the largest *HVR* and smallest *IGD* in all cases. As for the *C* metric, QHH-BS Pareto dominates the comparison MOEAs in all cases. Especially, QHH-BS fully Pareto dominates MOEA/D, MOABC, MOEO and MOGA in almost all cases, while MOEA/D, MOABC, MOEO and MOGA hardly Pareto dominates QHH-BS on almost all cases. All these indicate that the proposed QHH-BS achieves a better approximation to the true Pareto frontier.

To investigate the statistical significance, the widely used non-parametric Friedman test [45,46] is employed. Table 7 shows the statistical results derived from the Friedman test. Clearly, QHH-BS obtains the minimum rank in *IGD* and the maximum rank in *HVR* with a *p*-value lower than 0.0001. From all above, it is safe to conclude that QHH-BS significantly outperforms the involved comparison algorithms.

In addition, to provide better observation of the Pareto frontier solutions achieved by these algorithms, the Pareto frontiers of three cases with different scales are presented in Fig. 10. Note that, the termination criterion for these algorithms is the same. Compared with the comparison MOEAs, QHH-BS yields more well-distributed non-dominated solutions, and most of these solutions dominate those by the other state-of-the-art MOEAs. The reason can be attributed to the better exploration and exploitation achieved by Q-learning-based hyper-heuristic, and the good diversity and convergence provided by Bi-criterion selection.

## 6. Conclusion and future work

This work addresses an energy-aware mixed shop scheduling problem where speed-scaling policy and no-idle time strategy are adopted to reduce energy consumption and improve production efficiency. A MILP model is formulated to determine the sequence of job-shop and flow-shop products and the speed level of operations. Meanwhile, a multi-objective Q-learning-based hyper-heuristic with Bi-criteria selection (QHH-BS) is developed to solve this problem efficiently. Specifically, a problem-oriented three-layer encoding is devised, and a forward and backward decoding mechanism is developed to translate the above encoding into a feasible production plan. Pareto-based and indicator-based criteria are both utilized in the selection step to improve diversity and convergence. Q-learning with a multi-objective metric-based reward mechanism is applied to effectively and adaptively select an optimizer from {grey wolf operator, Jaya operator and crossover operator} in each iteration for better exploration and exploitation. Finally, three comparison experiments draw three conclusions:

(1) The proposed Bi-criteria selection is proved to be superior to two single-criterion selections due to the combination of strengths of the Pareto-based and indicator-based selection criteria.
(2) Q-learning-based hyper-heuristic shows the best performance since it combines the advantages of multiple low-level optimizers and adaptively selects an appropriate optimizer in each step in the search process.
(3) QHH-BS is compared with six state-of-the-art multi-objective algorithms including NSGA-II, MOEA/D, MOABC, MOEO, MOICA and MOGA. The results demonstrate that QHH-BS shows great superiority over other state-of-the-art algorithms in terms of convergence and diversity.

In future work, multiple batches of flow-shop products, setup times and transportation energy consumption should be considered to improve the applicability of the model. Besides, an effective approach is in need to reduce the number of calibrated parameters in the hyper-heuristics.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## CRediT authorship contribution statement

**Lixin Cheng:** Writing – original draft, Methodology, Software, Validation. **Qiuhua Tang:** Conceptualization, Writing – review & editing, Supervision. **Liping Zhang:** Writing – review & editing, Supervision. **Zikai Zhang:** Writing – review & editing, Supervision.

## Acknowledgment

This work is supported by the National Natural Science Foundation of China (No. 51875421, No. 51875420).

## Reference

[1] X. Shao, What is the right production strategy for horizontally differentiated product: standardization or mass customization, Int. J. Prod. Econ. 223 (2020) 1–11.
[2] J. Li, A. Wang, C. Tang, Production planning in virtual cell of reconfiguration manufacturing system using genetic algorithm, Int. J. Adv. Manuf. Technol. 74 (2014) 47–64.
[3] G. Wang, L. Gao, X. Li, P. Li, M.F. Tasgetiren, Energy-efficient distributed permutation flow shop scheduling problem using a multi-objective whale swarm algorithm, Swarm Evolut. Comput. 57 (2020) 1–12.
[4] C. Wang, R. Xu, J. Qiu, X. Zhang, AdaBoost-inspired multi-operator ensemble strategy for multi-objective evolutionary algorithms, Neurocomputing 384 (2019) 243–255.
[5] E.K. Burke, M. Gendreau, M. Hyde, G. Kendall, G. Ochoa, E. Ozcan, et al., Hyper-heuristics: a survey of the state of the art, J. Oper. Res. Soc. 64 (12) (2013) 1695–1724.
[6] N.V. Shakhlevich, Y.N. Sotskov, F. Werner, Complexity of mixed shop scheduling problems: a survey, Eur. J. Oper. Res. 120 (2) (2000) 343–351.
[7] L. Liu, Y. Chen, J. Dong, R. Goebel, G. Lin, Y. Luo, et al., Approximation algorithms for the three-machine proportionate mixed shop scheduling, Theor. Comput. Sci. 803 (2018) 57–70.
[8] A. Dugarzhapov, A. Kononov, A polynomial-time algorithm for the preemptive mixed-shop problem with two unit operations per job, J. Schedul. 19 (1) (2016) 1–12.
[9] M. Dai, D. Tang, A. Giret, M.A. Salido, Multi-objective optimization for energy-efficient flexible job shop scheduling problem with transportation constraints, Robot. Comput. Integr. Manuf. 59 (2019) 143–157.
[10] L. Meng, C. Zhang, X. Shao, Y. Ren, MILP models for energy-aware flexible job shop scheduling problem, J. Clean. Prod. 210 (2) (2019) 710–723.
[11] I. González-Rodríguez, J. Puente, J.J. Palacios, C.R. Vela, Multi-objective evolutionary algorithm for solving energy-aware fuzzy job shop problems, Soft Comput. 24 (2) (2020) 16291–16302.
[12] X. Wu, A. Che, A memetic differential evolution algorithm for energy-efficient parallel machine scheduling, Omega - Int. J. Manag. Sci. 82 (2019) 155–165.
[13] J. Chen, L. Wang, Z. Peng, A collaborative optimization algorithm for energy-efficient multi-objective distributed no-idle flow-shop scheduling, Swarm Evol. Comput. 50 (2019) 1–12.
[14] M. Abedi, R. Chiong, N. Noman, R. Zhang, A multi-population, multi-objective memetic algorithm for energy-efficient job-shop scheduling with deteriorating machines, Expert Syst. Appl. 157 (157) (2020) 1–17.
[15] A. Hasani, S.M.H. Hosseini, A bi-objective flexible flow shop scheduling problem with machine-dependent processing stages: trade-off between production costs and energy consumption, Appl. Math. Comput. 386 (2020) 1–23.
[16] F. Li, R. Cheng, J. Liu, Y. Jin, A two-stage R2 indicator based evolutionary algorithm for many-objective optimization, Appl. Soft Comput. 67 (2018) 245–260.
[17] E. Jiang, L. Wang, Z. Peng, Solving energy-efficient distributed job shop scheduling via multi-objective evolutionary algorithm with decomposition, Swarm Evol. Comput. (2020) 1–16.
[18] M. Li, S. Yang, X. Liu, Pareto or non-pareto: Bi-criterion evolution in multi-objective optimization, IEEE Trans. Evol. Comput. 20 (5) (2016) 645–665.
[19] H. Wang, L. Jiao, X. Yao, Two_Arch2: an improved two-archive algorithm for many-objective optimization, IEEE Trans. Evol. Comput. 19 (4) (2015) 524–541.
[20] X. Cai, Y. Li, Z. Fan, Q. Zhang, An external archive guided multiobjective evolutionary algorithm based on decomposition for combinatorial optimization, IEEE Trans. Evol. Comput. 19 (4) (2015) 508–523.
[21] J. Chen, L. Wang, Z. Peng, A collaborative optimization algorithm for energy-efficient multi-objective distributed no-idle flow-shop scheduling, Swarm Evol. Comput. 50 (2019) 30–42.
[22] R.H. Caldeira, A. Gnanavelbabu, A Pareto based discrete Jaya algorithm for multi-objective flexible job shop scheduling problem, Expert Syst. Appl. 50 (9) (2021) 30–42.
[23] J.H. Drake, A. Kheiri, E. Zcan, E.K. Burke, Recent advances in selection hyper-heuristics, Eur. J. Oper. Res. 285 (2) (2020) 405–428.
[24] J. Lin, Backtracking search based hyper-heuristic for the flexible job-shop scheduling problem with fuzzy processing time, Eng. Appl. Artif. Intell. 77 (1) (2019) 186–196.
[25] J. Lin, L. Zhu, K. Gao, A genetic programming hyper-heuristic approach for the multi--skill resource constrained project scheduling problem, Expert Syst. Appl. 140 (2019) 1–14.
[26] C. Qu, W. Gai, M. Zhong, J. Zhang, A novel reinforcement learning based grey wolf optimizer algorithm for unmanned aerial vehicles (UAVs) path planning, Appl. Soft Comput. 89 (2020) 1–15.
[27] I. Glcük, F.B. Ozsoydan, Q-learning and hyper-heuristic based algorithm recommendation for changing environments, Eng. Appl. Artif. Intell. 102 (2021) 1–18.
[28] H. Zhang, Y. Wu, R. Pan, G. Xu, Two-stage parallel speed-scaling machine scheduling under time-of-use tariffs, J. Intell. Manuf. 1 (2) (2020) 1–22.
[29] M.K. Marichelvam, M. Geetha, Ö. Tosun, An improved particle swarm optimization algorithm to solve hybrid flowshop scheduling problems with the effect of human factors – a case study, Comput. Oper. Res. 114 (2020) 104812–104821.
[30] S. Luo, L. Zhang, Y. Fan, Energy-efficient scheduling for multi-objective flexible job shops with variable processing speeds by grey wolf optimization, J. Clean. Prod. 234 (2019) 1365–1384.
[31] J. Li, J. Deng, C. Li, Y. Han, T. Jie, b. Zhang, et al., An improved Jaya algorithm for solving the flexible job shop scheduling problem with transportation and setup times, Knowl.-Based Syst. 200 (3) (2020) 1–13.
[32] R. Zhang, R. Chiong, Solving the energy-efficient job shop scheduling problem: a multi-objective genetic algorithm with enhanced local search for minimizing the total weighted tardiness and total energy consumption, J. Clean. Prod. 112 (1) (2016) 3361–3375.
[33] Y. Yefeng, y. Bo, W. Shilong, J. Tianguo, L. Shi, An enhanced multi-objective grey wolf optimizer for service composition in cloud manufacturing, Appl. Soft Comput. 87 (2020) 1–11.
[34] J. Li, H. Sang, Y. Han, C. Wang, K. Gao, Efficient multi-objective optimization algorithm for hybrid flow shop scheduling problems with setup energy consumptions, J. Clean. Prod. 181 (4) (2018) 584–598.
[35] E. Zitzler, L. Thiele, Multiobjective evolutionary algorithms: a comparative case study andt he strength pareto approach, IEEE Trans. Evol. Comput. 3 (4) (1999) 257–271.
[36] A.K. Shukla, R. Nath, P.K. Muhuri, Q.M.D. Lohani, Energy efficient multi-objective scheduling of tasks with interval type-2 fuzzy timing constraints in an Industry 4.0 ecosystem, Eng. Appl. Artif. Intell. 87 (2020) 103257–103274.
[37] Y. Liu, J. Liu, T. Li, Q. Li, An R2 indicator and weight vector-based evolutionary algorithm for multi-objective optimization, Soft Comput. 24 (7) (2020) 5079–5100.
[38] S. Jiang, J. Zhang, Y.S. Ong, A.N. Zhang, P.S. Tan, A simple and fast hypervolume indicator-based multiobjective evolutionary algorithm, IEEE Trans. Cybern. 45 (10) (2017) 2202–2213.
[39] P.A.N. Bosman, D. Thierens, The balance between proximity and diversity in multi--objective evolutionary algorithms, IEEE Trans. Evol. Comput. 7 (2) (2002) 174–188.
[40] R.H. Caldeira, A. Gnanavelbabu, T. Vaidyanathan, An effective backtracking search algorithm for multi-objective flexible job shop scheduling considering new job arrivals and energy consumption, Comput. Ind. Eng. 149 (2020) 1–27.
[41] Y. Li, W. Huang, R. Wu, K. Guo, An improved artificial bee colony algorithm for solving multi-objective low-carbon flexible job shop scheduling problem, Appl. Soft Comput. 95 (2020) 40–54.
[42] G.Q. Zeng, X.Q. Xie, M.R. Chen, J. Weng, Adaptive population extremal optimization-based PID neural network for multivariable nonlinear control systems, Swarm Evol. Comput. 44 (2019) 320–334.
[43] K.Z. Gao, Z. Cao, L. Zhang, Z. Chen, Q. Pan, A review on swarm intelligence and evolutionary algorithms for solving flexible job shop scheduling problems, IEEE/CAA J. Autom. Sin. 6 (4) (2019) 904–916.

[44] L. Pan, L. Li, R. Cheng, C. He, K.C. Tan, Manifold learning inspired mating restriction for evolutionary multi-objective optimization with complicated pareto sets, IEEE Trans. Cybern. 51 (2019) 3325–3337.
[45] L. Li, K. Lu, G. Zeng, L. Wu, M. Chen, A novel real-coded population-based extremal optimization algorithm with polynomial mutation: a non-parametric statistical study on continuous optimization problems, Neurocomputing 174 (2016) 577–587.
[46] D. Joaquín, G. Salvador, M. Daniel, H. Francisco, A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms, Swarm Evol. Comput. 1 (1) (2011) 3–18.