An Estimation of Distribution Algorithm for Mixed-Variable Newsvendor Problems

Feng Wang, *Member, IEEE*, Yixuan Li, Aimin Zhou, *Senior Member, IEEE*,
and Ke Tang, *Senior Member, IEEE*

Abstract—As one of the classical problems in the economic market, the newsvendor problem aims to make maximal profit by determining the optimal order quantity of products. However, the previous newsvendor models assume that the selling price of a product is a predefined constant and only regard the order quantity as a decision variable, which may result in an unreasonable investment decision. In this article, a new newsvendor model is first proposed, which involves of both order quantity and selling price as decision variables. In this way, the newsvendor problem is reformulated as a mixed-variable nonlinear programming problem, rather than an integer linear programming problem as in previous investigations. In order to solve the mixed-variable newsvendor problem, a histogram model-based estimation of distribution algorithm (EDA) called EDA$_{mvn}$ is developed, in which an adaptive-width histogram model is used to deal with the continuous variables and a learning-based histogram model is applied to deal with the discrete variables. The performance of EDA$_{mvn}$ was assessed on a test suite with eight representative instances generated by the orthogonal experiment design method and a real-world instance generated from real market data of Alibaba. The experimental results show that, EDA$_{mvn}$ outperforms not only the state-of-the-art mixed-variable evolutionary algorithms, but also a commercial software, i.e., Lingo.

Index Terms—Estimation of distribution algorithm (EDA), histogram model, mixed-variable optimization problem, newsvendor problem, orthogonal experiment design.

I. INTRODUCTION

TO ORDER how many products at what price, so as to make profit without taking excess risks, is one of the most elementary question for investors in economic market [1].

F. Wang and Y. Li are with the School of Computer Science, Wuhan University, Hubei 430072, China (e-mail: fengwang@whu.edu.cn; liyixuan@whu.edu.cn).

A. Zhou is with the School of Computer Science and Technology, East China Normal University, Shanghai 200062, China (e-mail: amzhou@cs.ecnu.edu.cn).

K. Tang is with the Shenzhen Key Laboratory of Computational Intelligence, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: tangk3@sustech.edu.cn).

In the scenarios like financial trading, marketing and product management [2]–[4], this question is typically formulated as a newsvendor model, in which the investors (e.g., retailers) make a decision on the order quantity of products based on the customers' actual (probably uncertain) demand of products. In other words, with a newsvendor model, a decision maker aims at seeking the optimal ordering strategy that leads to the maximum profit. Since its birth in 1950s, the newsvendor model has always been a critical research topic in economics, and has evolved with the development of new business models and new findings in economics, resulting in variants considering model parameters and decision variables [5], variants employing different objective functions [6], [7], and variants having various constraints [8]–[10].

Recently, due to the rise of new trading modes, such as presale [11] and crowdfunding [12], under the global market environments, two market factors, i.e., selling price and holding cost, which were more or less overlooked in previous research on newsvendor models, are gaining more and more attentions.

1) *Selling Price:* Previous newsvendor models treat selling price and demand of a product as two independent variables. However, with the development of social e-commerce, some new trading modes have emerged in trading market. It has been revealed that the selling price and actual demand are highly correlated in these emerging modes [13].

2) *Holding Cost:* With the expansion of retail companies, holding cost, i.e., the cost for warehousing products before sale, has become a major cost for many global retail companies. For example, it has been reported that the annual holding cost for Walmart is approximately 25% of its products' value [14]. However, most previous newsvendor models only concern purchasing cost, while assuming the holding cost is negligible.

By taking the above two emerging issues into account, this article proposes a new newsvendor model. Briefly speaking, while previous models assume the selling price of a product as a predefined constant and require seeking the optimal order quantity of the product, the proposed model also involves the selling price as a decision variable and requires a simultaneous optimization of both selling price and order quantity. Although such a modification appears to be minor at the first glance, the resultant optimization problem becomes a mixed integer nonlinear programming (MINLP) problem, rather than an integer linear programming (ILP) problem as

done in previous models, and thus brings significant challenges in the problem-solving aspect [15]–[17]. Hence, to address the MINLP problem induced by the new mixed-variable newsvendor model, a novel estimation of distribution algorithm (EDA) for mixed-variable newsvendor problem, denoted as EDA$_{mvn}$, is further proposed in this article. Specifically, EDA$_{mvn}$ employs a histogram model-based framework to estimate the probability distributions of different variables and sample new trail solutions. In EDA$_{mvn}$, an adaptive-width histogram (AWH) model and a learning-based histogram (LBH) model are proposed to accelerate the search for continuous and discrete variables, respectively. A tailored initialization scheme is also designed to ensure the initial solution to be feasible.

To summarize, the major contributions of this article include the following.

1) A newsvendor model is built, which considers both selling price and order quantity as decision variables and the holding cost as a constraint. The model transforms the newsvendor problem into a constrained MINLP problem. Given the fact that such a model represents the emerging scenarios of the newsvendor problems, tackling it will lead to significant applications in the economic market.

2) An EDA$_{mvn}$ algorithm is proposed for the new newsvendor model. Specifically, an AWH model is used to deal with the continuous variables, as well as an LBH model is used to deal with the discrete variables. Empirical studies show that EDA$_{mvn}$ not only outperforms some state-of-the-art evolutionary algorithms (EAs) for mixed-variable optimization problems, but also achieves better results than a commercial software, i.e., Lingo.

3) A benchmark test suite, which contains a synthetic benchmark suite and a real-world benchmark, is provided. The data of the real-world benchmark is from Alibaba, which can help us assess the performance of our model and algorithm in a real-market scenario. The synthetic benchmark suite considers various number of products (i.e., dimensionality of the problem), sensitivity of the demand on selling price (i.e., correlation between variables), and tightness of the constraints, is developed with the orthogonal experiment design approach. Such a synthetic test suite, originated from a real-world problem, provides a representative of the instance space of the newsvendor problem, and thus could be employed by interested readers in the evolutionary computation community for further study on MINLP in general.

The rest of this article is organized as follows. Section II briefly introduces the basic newsvendor model and gives detailed formalization of the new mixed-variable newsvendor problem. Section III describes the proposed EDA$_{mvn}$ for the mixed-variable newsvendor problem. Section IV introduces the detail procedures of select the representative test problems. The experimental studies are present in Section V to verify the performance of the new model and the proposed algorithm. Finally, Section VI concludes this article and highlights some directions for future work.

II. NEWSVENDOR MODELS AND SOLVERS

This section introduces the related assumptions and notations, the classical newsvendor model, our new model, and some existing solvers for newsvendor problems.

A. Assumptions and Notations

In the basic newsvendor model, a decision maker determines the order quantity of products to acquire as much profit as possible given some constraints. Here are some basic assumptions that are made to develop the mathematical model.

1) There is a single period to order the products.
2) The demand for the product is a random variable.
3) The ordering cost for the same product is fixed during the purchasing process.
4) The selling price is fixed without any discount during the selling period.
5) Excess ordering products that are more than the realized demand are sold with salvage value at the end of the selling period.
6) If the total demand is larger than the number of product, the customers' demand is responded by the actually ordering quantity.
7) The total available budget is limited.

In order to describe the model clearly, the following notations are defined.

1) M: The number of products.
2) m: The index for the mth product ($m = 1, \ldots, M$).
3) X: The order quantity vector for all products.
4) P: The selling price vector for all products.
5) O: The maximum total ordering cost in one investment.
6) H: The maximum total holding cost in one investment.
7) B: The maximum budget constraint.
8) x^m: The order quantity for the mth product.
9) d^m: The demand for the mth product.
10) c^m: The ordering cost of the mth product per unit.
11) p^m: The selling price of the mth product per unit.
12) s^m: The salvage value of the mth product per unit.
13) u^m: The understock cost of the mth product per unit.
14) h^m: The holding cost of the mth product per unit.

B. Basic Newsvendor Model

Let $(*)^+ = \max\{*, 0\}$, then the multiproduct newsvendor problem with a budget constraint can be expressed as follows:

$$\max \quad U(X) = \sum_{m=1}^{M} \Big[(p^m - c^m) \min\{x^m, d^m\} + (x^m - d^m)^+ s^m - (d^m - x^m)^+ u^m \Big]$$

$$\text{s.t.} \quad \sum_{m=1}^{M} c^m x^m \leq B \tag{1}$$

where $U(X)$ is the profit function, $(p^m - c^m) \min\{x^m, d^m\}$, $(x^m - d^m)^+ s^m$, and $(d^m - x^m)^+ s^m$ are the selling profit, the salvage residual value, and the understocking cost of the mth product, respectively.

It should be noted that it is not easy to acquire real demand in the future of the mth product. For convenience in mathematical derivation, the mth product's demand is always assumed to subject to a probability distribution, e.g., normal distribution or uniform distribution [7], [18]. But these probability distributions may not represent the real-market demand very well.

C. Mixed-Variable Newsvendor Model

With the development of the investment market, the basic newsvendor model can not describe newsvendor problems in the real-world investment market effectively. As mentioned above, most previous newsvendor models assume that the demand is subject to a fixed probability distribution, involve no more than two constraints, and use the order quantity as the only decision variable. The basic model does not consider the two new but important factors, i.e., selling price and holding cost. Specifically, in (1), the selling price is considered as a given constant. However, recent studies have found that it is related to the market demand [13]. In addition, the profit function in (1) ignores the products' holding cost.

For these reasons, we propose a new mixed-variable newsvendor model. This new model considers two kinds of decision variables: 1) order quantity and 2) selling price, and some crucial constraints simultaneously, i.e., total ordering cost constraint, total holding cost constraint, and price setting constraint.

1) Optimization Objective: The demand in real-market is always affected by some market factors, e.g., policy, time, and selling price. Among these factors, the selling price of product has crucial impact on the demand of product, which would directly affect the order quantity of the investors. However, the influence of the selling price on the market has always been neglected in the basic models. As the real demand of product dynamically changes with the selling price, investors may make improper decisions by using those old models. Some researchers have realized the importance of selling price on the demand and have discussed how to set appropriate price for each product in the newsvendor model [19], [20]. Therefore, both selling price and order quantity of products should be considered as decision variables.

The real demand of product changes with the selling price as follows [13]:

$$d^m = \alpha - \beta p^m \tag{2}$$

where α is the maximum demand of the mth product and β is the influence coefficient. This equation reveals that the demand of product d^m is not independent but price-deterministic, which means the demand of product will decrease as the selling price increases.

Here, the decision variables are $< X, P >$, which consist of discrete order quantity variables x and continuous selling price variables p. The non-negative linear additive demand broadly reflects that price is sensitive to investment market. In order to obtain a maximum profit, decision maker determines not only by the order quantity but selling price of products. The optimization objective function of the proposed model can be expressed as follows:

$$U(X, P) = \sum_{m=1}^{M} \Big[(p^m - c^m) \min\{x^m, d^m\} + (x^m - d^m)^+ s^m - (d^m - x^m)^+ u^m - h^m x^m \Big]$$

where $U(X, P)$ represents the profit of investment decision-making, and $h^m x^m$ is the holding cost of the mth product.

2) Constraints: In the real-world market, investors always make decisions with some restrained resources. In the new model, we consider the holding cost constraint as a resource constraint to make the model more realistic. What is more, it is unreasonable to set an infinite price. Therefore, the price setting constraint is also crucial. Since the total cost of an investment is limited, the sum of these two constraints should be subject to

$$O + H \le B$$

where

$$\sum_{m=1}^{M} c^m x^m \le O$$

$$\sum_{m=1}^{M} h^m x^m \le H.$$

These two resource constraints (ordering cost constraint and holding cost constraint) arose from real-world investments can describe the real market more reasonable and satisfactory.

Besides the above resource constraints, the decision-making interval of price should also be considered. As the demand will linearly decrease with price [see (2)], a smaller or larger price will make the demand be inappropriate for the real market. Hence, the price of product should be subject to a suitable interval, which can be described as follows:

$$\underline{p^m} \le p^m \le \overline{p^m}$$

where $\underline{p^m}$ and $\overline{p^m}$ are lower and upper price of mth product.

3) Mathematical Model: The new newsvendor model can be formulated as follows:

$$\max \ U(X, P) = \sum_{m=1}^{M} \Big[(p^m - c^m) \min\{x^m, d^m\} + (x^m - d^m)^+ s^m$$
$$- (d^m - x^m)^+ u^m - h^m x^m \Big]$$
$$\text{s.t.} \ \ O + H \le B$$
$$\sum_{m=1}^{M} c^m x^m \le O$$
$$\sum_{m=1}^{M} h^m x^m \le H$$
$$\underline{p^m} \le p^m \le \overline{p^m}. \tag{3}$$

From (3) and (1), we can see that the major difference between the new model and the basic model is on the decision variables and the constraints. In addition, the first term of the objective function [see (3)] becomes nonlinear since the demand of product depends on its price. And we can also find

that, for $m = 1, \ldots, M$, if $x^m > d^m$, the model is nonlinear, but degenerates into a linear model when $x^m < d^m$. Whereas, in the real market, we cannot guarantee that the order quantities of all products are smaller than their demands, so the new newsvendor problem is in general a nonlinear problem. As a result, this new model can be regarded as a model consisting of mixed decision variables (continuous price variables and discrete order quantity variables) and linear constraints.

D. Newsvendor Problem and MINLP Solvers

To solve newsvendor problems, some deterministic methods have been proposed. One of the most commonly used methods is the mathematical deduction method, e.g., in [21] and [22]. This algorithm is based on some strong assumptions that the objective function is continuous and differentiable and can easily get substantial gradient information. Another widely used deterministic method is the branch and bound method (BBM) [23]. BBM works well for simple problems, but for complex problem, it may fail to return a promising solution within a reasonable computational time. What is more, some other deterministic methods have also been proposed for newsvendor problems. Zhang et al. [24] proposed a binary solution method for the single-constraint model. Abdel-Malek and Areeratchakul [8] used a quadratic function to approximate the origin objective function, which simplifies the optimization problem but introduces error and can not get an accurate result for decision making.

To deal with MINLP problems, many heuristic methods have been proposed. Since EAs do not rely on the mathematical properties and are easy to be implemented, some EA-based methods have been developed to solve the MINLP problems. These methods mainly focus on the characteristics of the variables and different encoding schemes as well as evolutionary operators have also been proposed. MIES [25] is designed for continuous variables and nominal discrete and integer variables. GAMBIT [26] is proposed based on MIES and extends to solve constrained MINLP problems. A model-based EA that utilizes mixed-variable dependencies is proved to be successful for the black-box constraint problems [27]. Some other algorithms, which exploit the advantages of swarm intelligence algorithms have also been proposed for MINLP problems [28]–[32]. And some of these algorithms have been successfully applied in real-world applications, such as structural optimization problems in design industry [28], [33], [34], chemical engineering optimization problems [35], [36], scheduling problems in manufacturing [37], and insurance investment planning problems [38].

Unlike the previous MINLP problems, the mixed-variable newsvendor model proposed in this article is much more complex. First, as we can see in (3), the parameters p^m and d^m in the first term of the objective function have a tight correlation due to (2), and the new newsvendor problem is in general a nonlinear problem but degenerates into linear in some dimensions since the values of p^m and x^m are uncertain, which makes the landscape of the objective function become very rugged. Second, the number of product in this new model should be scalable in real applications, so the number of decision variables in the mixed-variable newsvendor model might be larger than the existing MINLP problems, which makes the decision space become more complex. As a result of this, it is very challengeable for the existing algorithms to find optimal solutions in the mixed-variable newsvendor model.

In the meanwhile, since the mixed-variable newsvendor problem can be formulated as a constrained MINLP model, there are several commercial MINLP solvers, e.g., Lingo[1] and CPLEX.[2] However, these commercial solvers are also very time-consuming, especially for complex large-scale problems.

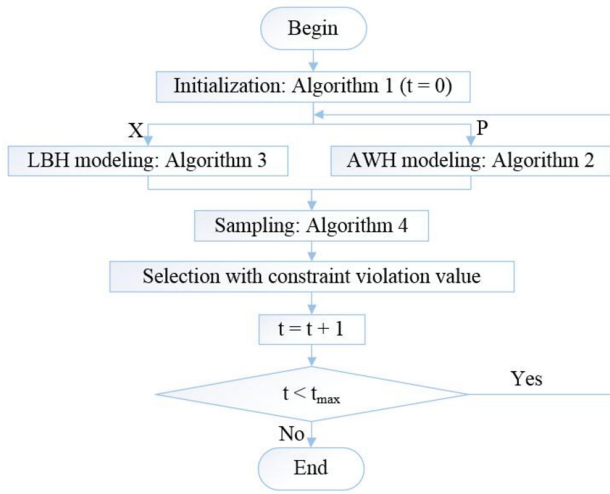## III. ESTIMATION OF DISTRIBUTION ALGORITHM FOR MIXED-VARIABLE NEWSVENDOR PROBLEMS

As mentioned in Section II-C, in the proposed mixed-variable newsvendor model, the objective function is discontinuous and nonlinear, and the constraints among the variables are also specific. Since EA is not restricted by the characteristics of the objective function, it might be a promising way to solve this new model. To the best of our knowledge, there is little work solving the newsvendor problems by using EAs. Although there are some EAs have been designed for newsvendor model [39], [40], the newsvendor models proposed in [39] and [40] are bi-objective, and only continuous variables are considered in [40]. Since these models are totally different from our single objective mixed-variable newsvendor model with multiple constraints, the corresponding algorithms are not suitable for solving this new mixed-variable model. As a result of this, it is required to design a new algorithm for the new mixed-variable newsvendor model.

EDA is a simple but effective EA. Unlike other EAs, EDAs learns the landscape of the search space by using a probabilistic model and generate new solutions by sampling from the model [41]–[45], and it has been widely applied to many real-world problems [46], [47]. Due to the continuous variables (the selling price of product p), the discrete variables (the order quantity of products x) and multiple constraints, the search space for the mixed-variable newsvendor problem is extremely complicated.

In this article, we propose a new EDA, called EDA_{mvn}, to deal with the mixed-variable newsvendor model. Specifically, a modified univariate histogram marginal model is used to deal with the continuous variables, as well as an LBH model is used to deal with the discrete variables. Besides, to improve the quality of the solutions in the initial population, a constraint handling method is embedded into EDA_{mvn}. And a constraints-based selection criterion is utilized in EDA_{mvn} to select individuals under the multiple constraints. The flowchart of EDA_{mvn} is shown in Fig. 1, and the parameter t is the current iteration, which increases after each generation. First, several individuals are initialized by a tailored initialization scheme. Then two histogram models, i.e., AWH model and LBH model, are used to deal with continuous variables P (selling price) and discrete variables X (order quantity), respectively. New individuals can be sampled from the probability

---

[1]https://www.lindo.com/index.php/products/lingo-and-optimization-modeling

[2]https://www.ibm.com/analytics/cplex-optimizer

Fig. 1. Flowchart of EDA$_{mvn}$.

Algorithm 1 $pop \leftarrow initialize(N)$

1: **Input:** Population size N;
2: **Output:** Initial population pop;
3: **for** i $\leftarrow$ 1 to N **do**
4:    **for** m $\leftarrow$ 1 **to** M **do**
5:       $p_i^m \leftarrow$ Uniform($\underline{p_i^m}, \overline{p_i^m}$);
6:       Estimate d_i^m by $\overline{(2)}$;
7:    **end for**
8:    $T \leftarrow \{1, 2, \ldots, M\}$;
9:    **while** $T \neq \emptyset$ **do**
10:      Randomly choose t from T;
11:      $x_i^t \leftarrow$ Uniform_int(0, $\overline{x_i^t}$);
12:      **if** $x_i^t > d_i^t$ **then**
13:        $x_i^t \leftarrow$ Gaussian(d_i^t, σ^2);
14:      **end if**
15:      **if** pop_i satisfies all the constraints **then**
16:        $T \leftarrow T\backslash\{t\}$;
17:      **else**
18:        $x_i^t \leftarrow 0$;
19:      **end if**
20:    **end while**
21:    Evaluate the fitness of the individual;
22: **end for**

---

models established by the AWH model and the LBH model and elite solutions can be selected by the real objective value and the constraint violation (CV). The algorithm will be terminated until t reaches the maximum number of generation $t_{\max}$. The details of EDA$_{mvn}$ are presented in the following sections.

A. Initialization

Since the mixed-variable newsvendor problem is an MINLP problem, we use a vector $<X, P>$ to represent each individual in EDA$_{mvn}$.

As mentioned above, multiple constraints are considered in the mixed-variable newsvendor model. According to the model proposed in (3), each solution should satisfy the following constraints.

  1) The value of the mth product price, denoted as p^m, should be in the range of $[\underline{p^m}, \overline{p^m}]$.
  2) The sum of the ordering cost cannot exceed the total ordering budget.
  3) The sum of the holding cost cannot exceed the maximum total holding cost.
  4) The total cost cannot exceed the given limit.

Among the above constraints, the first one is a box constraint and can be satisfied by sampling the p^m in a certain range. The others are resource constraints. Since each pair of p^m and x^m is sampled independently and uniformly, the solutions may exceed the constraints and introduce useless information into the search process. In order to promote search efficiency, it is better to initialize the search with feasible solutions. For this reason, an initialization process with constraint handling method is proposed and shown in Algorithm 1.

In the initialization process of an individual $<X, P>$, the price of each product is initialized in sequence. Once the price is decided in line 5, the real demand of each product can be estimated by (2) in line 6. Then, the order quantity of each product should be initialized in lines 10–14. For each individual, to ensure that the resource constraints are satisfied, after each order quantity is generated, the individual should

be checked by the constraints in line 15. Therefore, the later generated order quantities will be influenced by the former generated ones. In order to avoid that only the first several products are purchased for all individuals, the sequence to generate the product order quantity is random for each individual. By this initialization method, all individuals in the initial population are feasible and maintain a good diversity.

It should be noticed that the investor in our model is rational, therefore, the order quantity for a certain product will not be far from its real demand. This preference information is introduced in the initialization process as the prior information. Once a new generated order quantity is larger than the real demand, we repair it to a rational value. Here, we use the Gaussian distribution to generate this rational value in line 13 where σ denotes the investor's rationality degree and set as 0.1 in this article.

B. Mixed-Variable Probabilistic Models Building and Sampling

In EDA$_{mvn}$, a histogram model is used to solve the mixed-variable newsvendor problem and find the optimal solutions. The histogram model is a nonparametric method to estimate the probability distribution. In our approach, two different histogram models are used to deal with the continuous and discrete variables simultaneously. A modified univariate histogram marginal model, i.e., AWH model, is used to deal with the continuous variables, and an LBH model is used to deal with the discrete variables. In this section, we mainly introduce how to establish and update these two models and how to generate new individuals by sampling from these models.

1) Adaptive-Width Histogram Model: Inspired by [48], we propose an AWH model to modeling the population distribution of continuous variables in EDA$_{mvn}$. Both AWH model and VWH model [48] are histogram-based models with bins of variable width, but the methodologies of estimating the probability distribution of variables are different. Since the problem scale of the newsvendor model is larger than the benchmark test suite in [48], there is no extra count added when calculating the probability of each histogram. As a result, in order to improve the estimation accuracy of probability distribution in high-dimensional search space, in our AWH model, the probability distribution is updated by counting the individuals in each histogram. In addition, in order to improve the efficiency of the EDA$_{mvn}$, an unified framework exploring histogram-based models for mixed variables is developed, where the AWH model is only used to deal with continuous variables.

Let W represent the number of bins in the promising search area, in the AWH model, the search range of a continuous variable is divided into $(W+2)$ bins. And based on the distribution of the individuals in these bins, a probability model PRO$_c$ for the continuous variables will be established. Therefore, for the mth continuous variable p^m, the wth bin can be described by a tuple $<p^{m,w}, p^{m,w+1}, \text{PRO}_c^{m,w}>$, where $p^{m,w}$ and $p^{m,w+1}$ are the boundaries of the bin and PRO$_c^{m,w}$ denotes the probability that p^m is from the interval $[p^{m,w}, p^{m,w+1})$. The AWH model has the following features.

1) The W bins in the promising area have the same width and contain all the individuals, while the two bins at both sides are empty.
2) With the convergence of population, this promising area will become smaller and smaller. Hence, the width of the W bins in the promising area will become smaller adaptively, while the two bins at both sides will become larger.
3) The adaptively shrinking bins in the promising area can make the algorithm focus on searching high-precision solutions, in the meanwhile, every region in the search space has a chance to be search. Therefore, the exploration and exploitation can be automatically balanced.

To build the AWH model, the search space $[\underline{p^m}, \overline{p^m}]$ of the mth variable is partitioned into $(W+2)$ bins: $[p^{m,w}, p^{m,w+1})$, $w = 0, \ldots, W$ and $[p^{m,W+1}, p^{m,W+2}]$, where $p^{m,0} = \underline{p^m}$ and $p^{m,W+2} = \overline{p^m}$. Let $p_{min,1}^m$ and $p_{min,2}^m$ be the first and second minimum values of price, respectively, and $p_{max,1}^m$ and $p_{max,2}^m$ be the first and second maximum values, respectively, of the mth variables in the current population. Then set

$$p^{m,1} = \max\left\{p_{min,1}^m - 0.5(p_{min,2}^m - p_{min,1}^m), \underline{p^m}\right\}$$
$$p^{m,W+1} = \min\left\{p_{max,1}^m + 0.5(p_{max,1}^m - p_{max,2}^m), \overline{p^m}\right\}. \quad (4)$$

In the AWH model, the W bins in the promising area, range from $[p^{m,1}, p^{m,W+1}]$, have the same width, that is

$$\text{bin\_width} = \frac{1}{M}\left(p^{m,W+1} - p^{m,1}\right). \quad (5)$$

Let $A^{m,w}$ denote the count of individuals for the mth variable in the wth bin. It is obvious that $A^{m,1} = A^{m,W+2} = 0$. In order
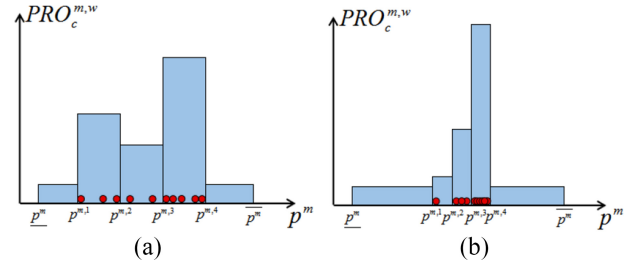

Fig. 2. Illustration of the search behavior of the AWH model ($W = 3$). (a) Early stage of the search process. (b) Later stage of the search process.

Algorithm 2 $PRO_c \leftarrow AWH\_model(pop, W)$

1: **Input:** the population *pop* and the number of bins W.
2: **Output:** the probability model PRO_c.
3: **for** m $\leftarrow$ 1 **to** M **do**
4:     Find $p_{min,1}^m, p_{min,2}^m, p_{max,1}^m$ and $p_{max,2}^m$;
5:     Calculate the boundaries of the promising area $p^{m,1}$ and $p^{m,W+1}$ by (4);
6:     Calculate the bin width by (5);
7:     **for** w $\leftarrow$ 1 **to** W+2 **do**
8:        Count the individuals which fall into the wth bin and set $A^{m,w}$ by (6);
9:     **end for**
10:     Establish the probability model PRO_c^m by (7);
11: **end for**

---

to avoid premature convergence, the count of individuals in the two unpromising search area which is out of $[p^{m,1}, p^{m,W+1}]$ will be assigned with a very small value ε to ensure that all search space of $[\underline{p^m}, \overline{p^m}]$ can be searched. Here, we update $A^{m,w}$ as

$$A^{m,w} = \begin{cases} A^{m,w} & \text{if } 2 \leq w \leq W+1 \\ \varepsilon & \text{if } w = 1, W+2, \text{ and } p^{m,w} > p^{m,w-1} \\ 0 & \text{if } w = 1, W+2, \text{ and } p^{m,w} = p^{m,w-1}. \end{cases}$$
$$(6)$$

After that, the probability model PRO$_c^m$ of p^m for the wth bin could be approximately estimated as

$$\text{PRO}_c^{m,w} = \frac{A^{m,w}}{\sum_{i=1}^{W+2} A^{m,i}}. \quad (7)$$

Fig. 2 shows the search behavior of the AWH model. The AWH model focuses on the promising search areas and is helpful to improve the performance in convergence. Especially in the later stage of the search process, the promising area becomes very small. In that case, the AWH model will spend more resources on exploitation. Meanwhile, the gradual changed promising area makes all the search space have a chance to be searched which can help the algorithm avoid getting trapped into local optima. The pseudo-code of the AWH model is shown in Algorithm 2.

2) Learning-Based Histogram Model: The LBH model is proposed for discrete variables, i.e., the order quantity x^m. By using the LBH model, a probability model PRO$_d^m$ for each available value of x^m is established.

In the LBH model, the probability for each available value v of x^m, $v \in \{0, 1, \ldots, \overline{x^m}\}$ is recorded as PRO$_d^{m,v}$. It should

---

Algorithm 3 $PRO_d \leftarrow LBH\_model(pop, t)$

1: **Input:** the population *pop* and the current generation *t*.
2: **Output:** the probability model PRO_d.
3: $\gamma \leftarrow \frac{t}{t_{max}}$;
4: **for** m $\leftarrow$ 1 **to** M **do**
5:    **for** v $\leftarrow$ 0 **to** $\overline{x^m}$ **do**
6:       count the individuals whose value of x^m equals to v
      and set as $count_v$;
7:       Update $PRO_d^{m,v}$ by (9);
8:    **end for**
9: **end for**

Algorithm 4 Offspring $\leftarrow$ Sampling(PRO_c, PRO_d)

1: **Input:** the probability models PRO_c and PRO_d.
2: **Output:** a new candidate solution offspring = <X,P>.
3: **for** m $\leftarrow$ 1 **to** M **do**
4:    $P_c \leftarrow$ Uniform(0,1);
5:    Select a bin w according PRO_c and P_c;
6:    $p^m \leftarrow$ Uniform$[p^{m,w-1}, p^{m,w})$;
7:    $P_d \leftarrow$ Uniform(0,1);
8:    Select an available value v from $\{0, 1, \ldots, \overline{x^m}\}$ according PRO_d and P_d;
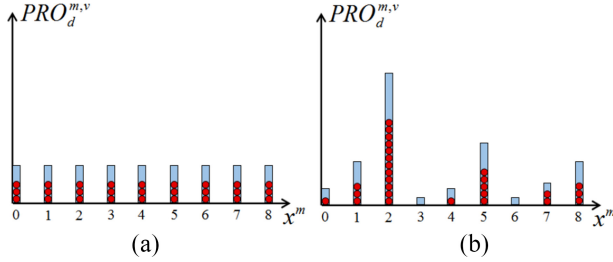9: $x^m \leftarrow v$;
10: **end for**

Fig. 3. Illustration of the search behavior of the LBH model ($\overline{x_m} = 8$). (a) Early stage of the search process. (b) Later stage of the search process.

be noticed that each available value is a normalized value. In the beginning, to ensure each available discrete value can be selected, the probability of each available value is set to be equal as

$$\mathrm{PRO}_d^{m,v}(0) = \frac{1}{x^m + 1}. \tag{8}$$

Then, the probability model will be updated by the following steps. In the generation *t*, first, count the individuals whose value of x^m equals to v and set it as $count_v$. Then update $\mathrm{PRO}_d^{m,v}$ as

$$\mathrm{PRO}_d^{m,v}(t) = (1-\gamma) \cdot \mathrm{PRO}_d^{m,v}(t-1) + \gamma \cdot \left(\frac{count_v}{N}\right) \tag{9}$$

where N is the size of population and t the current generation number. γ describes the learning rate of population. In this article, the parameter γ is set

$$\gamma = \frac{t}{t_{\max}} \tag{10}$$

where $t_{\max}$ is maximum number of generation and t is the current iteration. In this way, as the number of generations increases, γ will increase gradually, which makes the LBH model use more information from good quality individuals to update the probability model PRO_d^m.

Fig. 3 illustrates the search behavior of the LBH model ($\overline{x_m} = 8$). Fig. 3(a) and (b) shows the early stage ($\gamma = 0$) and the later stage ($\gamma = 1$) of the search process, respectively. At the beginning, the probability of each available value is set with same value. After some iterations, the solutions with some values (such as 2 and 5 in this example) are better than others, which makes these values have higher probability than others and become more likely to be sampled. The pseudo-code to update the LBH model is summarized in Algorithm 3.

3) Sampling: After the histogram models are established, the new individuals can be generated by sampling from the model. For a continuous variable p^m, a bin w is first selected by a randomly generated probability, and then p^m is uniformly sampled from $[p^{m,w-1}, p^{m,w})$. For a discrete variable x^m, an available value v from $\{0, 1, \ldots, \overline{x^m}\}$ is selected by a randomly generated probability. The procedure for generating new individual is shown in Algorithm 4.

C. Constraints-Based Selection

The penalty function method is a widely used method for handling constraints and producing selection pressure. EDA_{mvn} adopts a linear penalty function to select promising individuals.

By using the linear penalty function, the fitness values assigned to individuals can be divided into two parts, i.e., the real objective value and the CV

$$f(X, P) = U(X, P) + \mathrm{CV}(X, P) \tag{11}$$

where $U(X, P)$ is the real objective value of the obtained profit. The elite solutions will be selected according to the fitness value $f(X, P)$. The CV is calculated as follows:

$$\mathrm{CV}(X, P) = k_1 \sum_{i=1}^{c_1} |\min(g_i(X, P), 0)| + k_2 \sum_{j=1}^{c_2} |h_j(X, P)| \tag{12}$$

where $g_i(X, P)$ is the violation value of ith inequality constraint and $h_j(X, P)$ is the violation value of jth equality constraint, c_1 and c_2 are the number of inequality constraints and equality constraints, respectively, and k_1 and k_2 are the penalty parameters.

D. Algorithm Framework

Based on the methods presented above, the overall procedure of EDA_{mvn} can be presented in Algorithm 5. In line 6, an initial population is generated by the constraint handling method. In lines 9 and 10, the probability models for the continuous and discrete variables are generated, respectively. And then, in line 12, new individuals are generated by sampling from these two probability models. In addition, a global elite replacement strategy is used to select the best N solutions in each generation. Therefore, an *archive* is employed in our algorithm.

Algorithm 5 Framework of EDA_{mvn}

1: **Input:**
2: 　1) N: the population size.
3: 　2) t_{max}: the maximum number of generation.
4: 　3) W: the number of bins used in the AWH model.
5: **Output:** the best solution and its objective value.
6: $pop \leftarrow initialize(N)$;
7: **for** t $\leftarrow$ 1 **to** t_{max} **do**
8: 　archive $\leftarrow \emptyset$;
9: 　$PRO_c \leftarrow AWH\_model(pop, W)$;
10: 　$PRO_d \leftarrow LBH\_model(pop, t)$;
11: 　**for** i = 1 **to** N **do**
12: 　　offspring $\leftarrow$ Sampling(PRO_c, PRO_d);
13: 　　archive $\leftarrow$ archive $\bigcup$ offspring;
14: 　**end for**
15: 　archive $\leftarrow$ archive $\bigcup$ pop;
16: 　pop.clear();
17: 　Select the best N individuals according to (11) from archive and add them to pop;
18: **end for**

TABLE I
FACTORS AND LEVELS FOR ORTHOGONAL EXPERIMENT

Factor	level 1	level 2	level 3	level 4
M	20	40	80	100
$\overline{x^m}$	50	100	\	\
CF	Weak	Strong	\	\
PS	Yes	No	\	\

TABLE II
SYNTHETIC BENCHMARK SUITE FOR MIXED-VARIABLE NEWSVENDOR
PROBLEMS. THE NUMBERS IN PARENTHESES REPRESENT THE LEVEL OF
EACH VALUE CORRESPONDING TO TABLE I

Instance	Factors			
	M	$\overline{x^m}$	CF	PS
F_1	20(1)	50(1)	Weak(1)	Yes(1)
F_2	20(1)	100(2)	Strong(2)	No(2)
F_3	40(2)	50(1)	Weak(1)	No(2)
F_4	40(2)	100(2)	Strong(2)	Yes(1)
F_5	80(3)	50(1)	Strong(2)	No(2)
F_6	80(3)	100(2)	Weak(1)	Yes(1)
F_7	100(4)	50(1)	Strong(2)	Yes(1)
F_8	100(4)	100(2)	Weak(1)	No(2)

IV. BENCHMARK SUITE FOR THE MIXED-VARIABLE NEWSVENDOR PROBLEM

To the best of our knowledge, few benchmark instance for newsvendor problem could be found in the literature. To systematically evaluate the usefulness of the mixed-variable newsvendor model and the effectiveness of the proposed EDA_{mvn}, we first design a synthetic benchmark suite by using orthogonal experiment design method, and then use the real market data from Alibaba as a real-world benchmark.

A. Synthetic Benchmark

In the mixed-variable newsvendor model, there are four main features: 1) M; 2) $\overline{x^m}$; 3) constraint force (CF); and 4) price sensitivity (PS). Herein, M is the number of products, $\overline{x^m}$ is the maximum value of the discrete variables, CF is the constraints force, and PS is the sensitivity of demand to price. The meaning and the values of these factors are introduced as follows.

1) M determines the dimensions of the search space. In general, solving problem with low dimension is easier than that with high dimension [49], [50]. Therefore, to test the performance of algorithms for solving problems with different dimensions, the value of M is set as {20, 40, 80, 100}.

2) $\overline{x^m}$ determines the search scope of the discrete variables. To test the performance of discrete optimization methods for handling discrete variables with different search scope, $\overline{x^m}$ is set as 50 and 100, respectively.

3) CF determines the complexity of the search space. In order to analyze the influence of constraints on algorithm performance, two different kinds of CF are set in our benchmark suite, denote as *weak* and *strong*, respectively. In the mixed-variable newsvendor model, the CF is reflected on the values of total ordering cost and total holding cost. For fairness, these two values are correlated to M for different problems. If CF = Weak, these two values are set as 100 times of the value of M. Otherwise, they are set as 50 times of the value of M.

4) PS determines the correlation between the continuous variable (i.e., the selling price) and the discrete variable (i.e., the order quantity), and is used to validate the compatibility of the continuous optimization method and discrete optimization method. In the mixed-variable newsvendor model, the PS is denoted as the ratio between the parameter α and β in (2). To ensure the demand in the market is non-negative, the lower bound of α/β should be given by the selling price. In our test instances, in order to differentiate the difference of PS (PS = Yes or PS = No), we set different values for α/β. If PS = Yes, the ratio α/β is set as 8 for each product, otherwise, this ratio is set as 20.

To sum up, M is regarded as a factor at four different levels and $\overline{x^m}$, CF, and PS are regarded as factors at two different levels, as illustrated in Table I.

In order to cover all the feature space, we should test all combinations of these factors in traditional approaches and need at least $4 \times 2^3 = 32$ test instances, which is not necessary and economical. In this article, the orthogonal experiment design method, which can reduce the number of experiments but still achieve satisfactory results, is applied to develop a synthetic benchmark suite. Herein, based on the common used orthogonal factor table $L_8(4 \times 2^4)$, we design the orthogonal factor table for the mixed-variable newsvendor problems, as illustrated in Table II. Each row in Table II denotes to the parameter settings of a representative test instance in the synthetic benchmark suite.

The values of the parameters for each instance are set by considering different scenarios of the newsvendor problem. According to the previous research on newsvendor problems [39], [51], [52], for each test instance, the parameters of products are generated as follows and let $t \sim U(a, b)$

denote that t is uniformly generated from interval $[a, b]$: the holding cost $h^m \sim [0, 10]$, the ordering cost $c^m \sim [0, 10]$, and the understock cost $u^m \sim [0, 10]$, for $m = 1, 2, \ldots, M$. The detailed information of each instance including parameter setting is shown in the supplementary material.

B. Real-World Benchmark

The data of the real-world benchmark is from Alibaba.[3] This data set provides the information of 963 products, including the selling price, overstock cost, understock cost, etc. Since the original data cannot be directly used in the experiment, it is required to do some preprocessing work on the data set before our experimental study. The preprocessing procedure is introduced in the supplementary material and the processed data we used can be found at our website.[4] In order to test the usefulness of the mixed-variable newsvendor model and the performance of EDA$_{mvn}$, 50 products are randomly selected from the processed data set and used in our experiment. This real-world benchmark from real market data is denoted as *RMdata* in this article.

V. EXPERIMENTAL STUDIES

This section empirically studies the performance of EDA$_{mvn}$ on the given benchmark suite. Specifically, there are two major concerns that we have to validate by experiments.

The first one is the effectiveness of the proposed EDA$_{mvn}$ algorithm. In the mixed-variable newsvendor model, domain knowledge, such as PS, is employed to make the model more practical. Whether the domain knowledge can be used by the proposed EDA$_{mvn}$ algorithm to help the decision makers obtain higher profit is required to be verified at first. And it is also a baseline to roughly check the performance of EDA$_{mvn}$.

The second one is the efficiency of the proposed EDA$_{mvn}$ algorithm. To verify this, we compare our algorithm with some state-of-the-art algorithms or solvers. It should be pointed out that although there is no specific algorithm designed directly to solve this kind of newsvendor problem, there are several algorithms for MINLP problems or mixed-variable problems. For example, GA-PSO [31], ACO_{mv} [30], and DE_{mv} [53] are three different kinds of EAs for the mixed-variable problems, and the evolutionary mechanisms behind these three algorithms are very different to EDA$_{mvn}$. AEDA [38] is an EDA-based algorithm for mixed-variable optimization problems, in which histogram model is only used to handle discrete variables. Therefore, we select these four algorithms as comparison algorithms. In addition, a commercial solver Lingo, which can solve many kind of mathematical programming problems, is also selected as comparison in the experiment.

A. Parameter Settings

To have a fair comparison, the parameters of the comparison algorithms are set following the recommended setting in the

original paper. The parameter settings of the experiments are as follows.

1) All algorithms except Lingo are executed independently for 30 runs with 300 000 maximum function evaluations (FEs).
2) *EDA$_{mvn}$:* The number of bins used to establish the AWH model $W = 2000$; the population size $N = 600$; the penalty parameters $k_1 = 1000$ and $k_2 = 0$ because there is no equality constraint in our model.
3) *GA-PSO:* The population size $N = 300$; the probability of crossover $P_c = 1.0$; and the probability of mutation $P_m = 0.1$.
4) *ACO$_{mv}$:* The influence of best quality solutions $q = 0.05099$; the width of the search $\xi = 1$; and the population size $N = 20$.
5) *DE$_{mv}$:* The parameters for discrete variables $F_d = 0.2$ and $CR_d = 0.2$; the parameters for continuous variables $F_c = 0.5$ and $CR_c = 0.5$; the mutation strategy is "DE/current-to-pbest"; and the population size $N = 100$.
6) *AEDA:* The number of individuals selected for model estimation, NP_{best}, is set as 45% of the whole population; and the population size $N = 300$.

In addition, the other four comparison algorithms do not have mechanisms to deal with constraints. Since the mixed-variable newsvendor model involves multiple constraints. To have a fair comparison, the constraint-based selection criterion is also employed in these algorithms.

B. Validation of Algorithm Effectiveness

The main difference between the mixed-variable newsvendor model and the classical model is that we consider both product's selling price and order quality as decision variables and the real demand of product changes with price in the mixed-variable newsvendor model, while only order quality has been considered as decision variable in the classical model. Therefore, by using the mixed-variable newsvendor model to help us make decision, the price of each product is selected in the range $[\underline{p_m}, \overline{p_m}]$, while by using the classical model, we set each product's price as a fixed number $\overline{p_m}$. The reason why we set $\overline{p_m}$ as the fixed number is that when the decision makers use the classical model, they don't care about the relationship between the real demand and the price. Therefore, in order to get the maximum profit, they will set the price as higher as possible.

Each instance in Section IV represents a typical application scenario. And we use the mixed-variable newsvendor model and the classical model to help us make decisions in these scenarios, respectively. The optimization algorithm used here is EDA$_{mvn}$. Here, each instance is tested independently for 30 times with 300 000 FEs. The results on the benchmark suite are shown in Table III.

Although in instance F_5, the performance of EDA$_{mvn}$ on the mixed-variable newsvendor model shows a little worse than the classical model, the results in other instances show that EDA$_{mvn}$ can help investors get significantly higher profit on the mixed-variable newsvendor model, including in the real-world benchmark *RMdata*. These results illustrate that

TABLE III
COMPARISON RESULTS BETWEEN THE MIXED-VARIABLE NEWSVENDOR
MODEL AND CLASSICAL MODEL ON THE BENCHMARK SUITE. THE BEST
RESULTS ARE SHOWN IN BOLD TYPE. "+," "−," OR "≈" INDICATES THE
PERFORMANCE OF EDA$_{mvn}$ ON THE CLASSICAL MODEL IS BETTER
THAN, WORSE THAN OR SIMILAR TO THE MIXED-VARIABLE
NEWSVENDOR MODEL AT 0.95 SIGNIFICANCE LEVEL BY
A WILCOXON RANK SUM TEST

Instance		Mixed-variable Model	Classical Model	Improvement Ratio
F_1	MEAN	**1.572E+01**	1.347E+01$^-$	16.67%
	SD	**1.053E+00**	3.892E-14	
F_2	MEAN	**5.709E+02**	4.883E+02$^-$	16.93%
	SD	**7.159E+00**	1.245E-12	
F_3	MEAN	**1.906E+03**	1.632E+03$^-$	16.81%
	SD	**2.723E+01**	0.000E+00	
F_4	MEAN	**1.260E+02**	1.070E+02$^-$	17.70%
	SD	**4.519E+00**	7.784E-14	
F_5	MEAN	3.588E+03	**3.643E+03**$^+$	−1.50%
	SD	5.962E+01	**1.279E+00**	
F_6	MEAN	**2.931E+02**	1.985E+02$^-$	47.67%
	SD	**2.209E+01**	6.227E-13	
F_7	MEAN	**2.166E+02**	1.994E+02$^-$	8.63%
	SD	**3.538E+00**	0.000E+00	
F_8	MEAN	**2.158E+03**	1.886E+03$^-$	14.39%
	SD	**5.828E+01**	4.982E-12	
$RMdata$	MEAN	**3.076E+05**	2.575E+05$^-$	19.46%
	SD	**2.184E+03**	7.970E-10	

EDA$_{mvn}$ is effective to solve the mixed-variable newsvendor model.

C. Validation of Algorithm Efficiency

To verify the efficiency of EDA$_{mvn}$, we compare EDA$_{mvn}$ with other state-of-the-art algorithms, i.e., GA-PSO, AEDA, DE$_{mv}$, and ACO$_{mv}$, and a commercial solver Lingo on the benchmark suite designed in Section IV. The comparison results on the benchmark suite are detailed in Table IV. From these results, we can see that EDA$_{mvn}$ are significantly better than other algorithms in most of the test instances except on F_1 and F_4. On F_1, the result of EDA$_{mvn}$ is slightly worse than DE$_{mv}$ but better than the other algorithms. The performance of EDA$_{mvn}$ on the real-world benchmark $RMdata$ also outperforms others. These results validate that EDA$_{mvn}$ is more efficient than other algorithms to solve the mixed-variable newsvendor model.

In addition, to testify the effect of each factor in the synthetic benchmark suite on the algorithm performance, we evaluate the average performance of each algorithm on each factor. Detailed analyses of the results are illustrated as follows, and in order to take a fair comparison between different instances, all results on the synthetic benchmark suite are normalized finally.

1) Comparison Under Different Dimensions of Search Space (M): Table V shows the comparison results under different dimensions of the search space. All the results are shown by the normalized values. k_i^j is the mean value of the experimental results of factor j at level i in the synthetic benchmark suite. The information of factor level can be found in the above Table I.

From Table V, we can see that EDA$_{mvn}$ shows a great performance in the high-dimensional problems and a

promising results in the low-dimensional problems. In the meantime, AEDA also shows satisfying results in the high-dimensional problems. The major reason is that according to the increasing of dimension, the search space increases geometrically. While in the EDA framework, even at the later stage that the population converge into a very small area, all search space still has a chance to be searched. This feature makes EDA framework have a great global search ability and have a higher chance to find high-quality solutions in the high-dimensional problems.

In addition, although EDA$_{mvn}$ and AEDA are based on the EDA framework, EDA$_{mvn}$ still outperforms AEDA on these test instances. A possible reason is that, AEDA assumes the probability model as a given distribution like Gaussian distribution or Cauchy distribution, but EDA$_{mvn}$ uses two different histogram models to estimate the distribution which can represent the real distribution better. As the estimation of probability models for both discrete variables and continuous variables are histogram models, the reproduction mechanisms in these two histogram models for different types of decision variables are consistent and compatible.

2) Comparison Under Different Number of Available Values of Discrete Variables ($\overline{x^m}$): As illustrated in Table VI, EDA$_{mvn}$ shows a satisfying result in various search range, even though the number of available values is large. This result illustrates that the LBH model for the discrete variables works well.

In the LBH model, the adaptive learning-based strategy is employed, which can make the algorithm avoid early convergence and have a good balance between exploration and exploitation in the discrete search space. As the increasing number of available values make the combinations of the discrete variables grow vigorously, the effect of the LBH model is obviously greater than the other algorithms which utilize different methods for discrete variables, especially in the problems where the search scope of variables is extremely large.

3) Comparison Under Different Degree of Constraint Forces: The multiple constraints in the mixed-variable newsvendor model make the search space become more complicated, especially when the CF becomes strong. The results shown in Table VII also validate that EDA$_{mvn}$ has a great performance in these test instances, even if in the problems with strong CF. There are two reasons as follows. First, the AWH model makes every region in the continuous search space have a chance to be searched in each iteration. This feature can help EDA$_{mvn}$ avoid getting trapped into the local optima under the complicated environment. Second, the LBH model is also very suitable for the complex problems by using the learning-based strategy and the adaptive learning rate. Therefore, the combination of these two models can help the algorithm find better solutions.

4) Comparison Under Different Price Sensitivity: From Table VIII, we can see that EDA$_{mvn}$ and DE$_{mv}$ outperform other algorithms under this comparison. In some ways, the PS is the connection between the continuous and discrete variables. A high selling price will decrease the real demand and influence the order quantity, or vice verse. In EDA$_{mvn}$, two histogram models are used to deal with

TABLE IV
COMPARISON RESULTS FOR EDA$_{mvn}$, GA-PSO, ACO$_{mv}$, DE$_{mv}$, AEDA, AND LINGO ON THE BENCHMARK SUITE. THE BEST RESULTS ARE SHOWN IN
BOLD TYPE. "+," "−," OR "≈" INDICATES THE PERFORMANCE IS BETTER THAN, WORSE THAN OR SIMILAR TO THE PERFORMANCE OF
EDA$_{mvn}$ AT 0.95 SIGNIFICANCE LEVEL BY A WILCOXON RANK SUM TEST

Instance		EDA_{mvn}	GA-PSO	ACO_{mv}	DE_{mv}	AEDA	Lingo
F_1	MEAN	1.572E+01	1.556E+01$^-$	1.551E+01$^\approx$	**1.598E+01**$^+$	1.548E+01$^-$	1.539E+01$^-$
	SD	1.053E+00	1.401E+00	2.352E+00	**1.550E-01**	8.548E-01	\
	Normalized Value	0.555	0.284	0.205	**1.000**	0.156	0.000
F_2	MEAN	**5.709E+02**	5.685E+02$^-$	5.437E+02$^-$	5.693E+02$^-$	5.454E+02$^-$	5.385E+02$^-$
	SD	**7.159E+00**	2.159E+01	1.156E+02	1.310E+01	1.810E+01	\
	Normalized Value	**1.000**	0.924	0.158	0.952	0.213	0.000
F_3	MEAN	**1.906E+03**	1.775E+03$^-$	1.715E+03$^-$	1.838E+03$^-$	1.776E+03$^-$	1.654E+03$^-$
	SD	**2.723E+01**	2.062E+02	4.766E+02	6.656E+01	6.092E+01	\
	Normalized Value	**1.000**	0.479	0.243	0.729	0.482	0.000
F_4	MEAN	1.260E+02	1.271E+02$^+$	1.197E+02$^\approx$	**1.360E+02**$^+$	1.262E+02$^\approx$	1.211E+02$^-$
	SD	4.519E+00	8.103E+00	1.468E+02	**7.690E+00**	6.229E+00	\
	Normalized Value	0.386	0.452	0.000	**1.000**	0.398	0.090
F_5	MEAN	**3.588E+03**	2.823E+03$^-$	2.252E+03$^-$	3.175E+03$^-$	2.861E+03$^-$	3.132E+03$^-$
	SD	**5.962E+01**	1.111E+03	2.667E+03	3.248E+02	2.452E+02	\
	Normalized Value	**1.000**	0.427	0.000	0.691	0.456	0.659
F_6	MEAN	**2.931E+02**	2.289E+02$^-$	1.420E+02$^-$	2.783E+02$^-$	2.880E+02$^-$	1.902E+02$^-$
	SD	**2.209E+01**	9.418E+01	5.268E+02	3.188E+01	1.345E+01	\
	Normalized Value	**1.000**	0.575	0.000	0.902	0.967	0.319
F_7	MEAN	**2.166E+02**	1.881E+02$^-$	1.014E+02$^-$	2.127E+02$^-$	2.015E+02$^-$	1.675E+02$^-$
	SD	**3.538E+00**	8.057E+01	4.337E+02	8.882E+00	7.238E+00	\
	Normalized Value	**1.000**	0.753	0.000	0.967	0.869	0.574
F_8	MEAN	**2.158E+03**	1.658E+03$^-$	9.364E+02$^-$	1.963E+03$^-$	2.067E+03$^-$	1.836E+03$^-$
	SD	**5.828E+01**	4.620E+02	1.773E+03	1.003E+02	6.011E+01	\
	Normalized Value	**1.000**	0.591	0.000	0.841	0.926	0.737
$RMdata$	MEAN	**3.076E+05**	2.987E+05$^-$	2.303E+05$^-$	2.988E+05$^-$	2.926E+05$^-$	2.869E+05$^-$
	SD	**2.184E+03**	1.326E+04	1.697E+05	6.898E+03	8.095E+03	\

TABLE V
COMPARISON RESULTS ON THE SYNTHETIC BENCHMARK SUITE WITH DIFFERENT DIMENSIONS OF THE SEARCH SPACE. THE RESULTS SHOWN
BY THE NORMALIZED VALUES. EACH ROW SHOWS THE RESULTS OF DIFFERENT ALGORITHMS UNDER A CERTAIN FACTOR LEVEL.
THE BEST RESULTS ARE SHOWN IN BOLD TYPE

	M	EDA_{mvn}	GA-PSO	ACO_{mv}	DE_{mv}	AEDA	Lingo
k_1^1	20	0.778	0.604	0.181	**0.976**	0.185	0.000
k_2^1	40	0.693	0.465	0.121	**0.864**	0.440	0.045
k_3^1	80	**1.000**	0.501	0.000	0.797	0.711	0.489
k_4^1	100	**1.000**	0.672	0.000	0.904	0.897	0.656

TABLE VI
COMPARISON RESULTS ON THE SYNTHETIC BENCHMARK SUITE WITH DIFFERENT NUMBER OF AVAILABLE VALUES OF DISCRETE VARIABLES.
THE RESULTS SHOWN BY THE NORMALIZED VALUES. EACH ROW SHOWS THE RESULTS OF DIFFERENT ALGORITHMS UNDER A CERTAIN
FACTOR LEVEL. THE BEST RESULTS ARE SHOWN IN BOLD TYPE

	$Range$	EDA_{mvn}	GA-PSO	ACO_{mv}	DE_{mv}	AEDA	Lingo
k_1^2	[0,50]	**0.889**	0.486	0.112	0.847	0.491	0.308
k_2^2	[0,100]	0.846	0.636	0.040	**0.924**	0.626	0.286

the continuous and discrete variables, respectively, while in DE$_{mv}$, original DE and set-based DE are used. The search mechanisms behind each of these two algorithms are highly compatible and consistent. This feature makes these two algorithms have a better performance to deal with these problems where continuous variables are correlate with discrete ones [53]. The results in Table VIII show that EDA$_{mvn}$ and DE$_{mv}$ have the best performance and validate the compatible and consistent search mechanisms do have good efficiency.

From the above results and analyses, we can conclude that the proposed EDA$_{mvn}$ outperforms other algorithms, which can be attributed to the following three reasons. First, EDA has a great performance in global search capability, which makes it more suitable to solve the problems with complicated search space. Second, the histogram model-based framework, which uses the histogram model to establish the probability models for both continuous and discrete variables, has a great generalization capability. Last but no least, the adaptive scheme in both AWH model and LBH model can help the algorithm get good balance between exploration and exploitation, which makes the search process avoid from trapping into local optimum.

5) Convergence Process: Fig. 4 shows the convergence process of EDA$_{mvn}$ and other state-of-the-art algorithms on the benchmark suite. We can see that the convergence speed of EDA$_{mvn}$ is not outstanding on the simple problems (i.e., F_1–F_4), but it outperforms on the other complex problems (i.e., F_5–F_8), which is high-dimensional. EDA$_{mvn}$ also shows a faster average convergence speed on the real-world benchmark
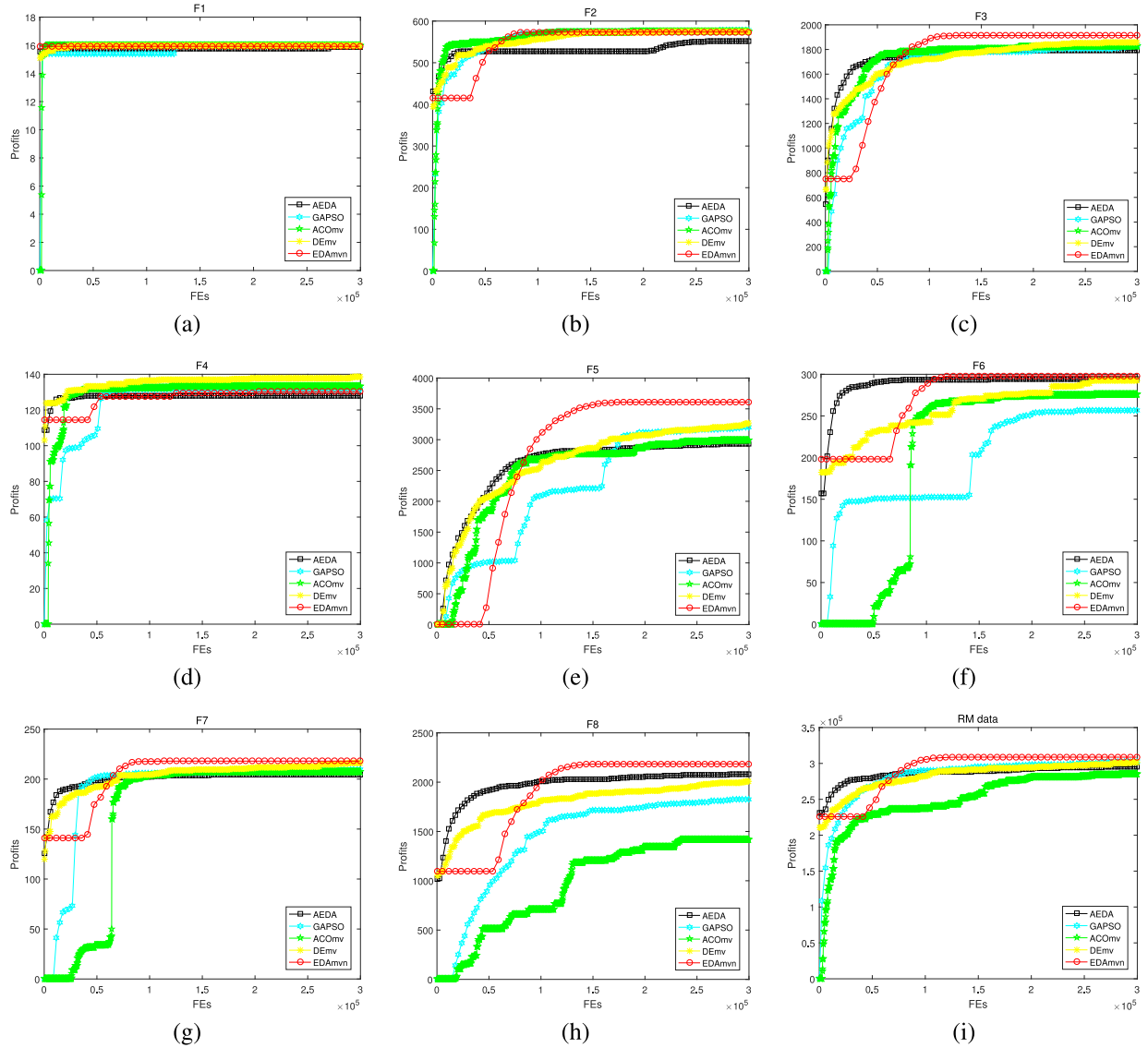
Fig. 4. Convergence process on the benchmark suite. Each curve shows the convergence process with the largest profit obtained by each algorithm out of all the 30 runs. (a)–(h) are the convergence process on F_1–F_8, respectively. (i) is the convergence process on RMdata.

RMdata, which illustrates that EDA$_{mvn}$ is effective in real applications. Moreover, AEDA shows a faster convergence speed than other algorithms on most of the test instances at the beginning of the search process, but the final solutions of it are not good. As for the other three algorithms, DE_{mv} shows better than GA-PSO and ACO$_{mv}$.

In addition, we can draw the following two conclusions from these figures. First, the initial population of EDA$_{mvn}$ shows a higher quality than others. This is because in the initialization of EDA$_{mvn}$, we use a constraint handling method to generate more feasible solutions. This method can enhance the solution quality. Second, in EDA$_{mvn}$, the fitness of the optimal solution keeps relative stable with a low value at the beginning of evolution process but increases very quickly after several iterations. The reason is that at the early stage, the promising area in the AWH model is still very large, while the adaptive width of histogram is helpful to shrink this promising area later. With the promising area becomes smaller and smaller,

the probability to find better solutions increases significantly, thus the convergence process of EDA$_{mvn}$ will speed up.

D. Sensitivity Analysis to Parameters in EDA$_{mvn}$

The population size N and the number of bins W are two crucial parameters to EDA$_{mvn}$. In this section, we discuss the effect of these two parameters to EDA$_{mvn}$ based on the benchmark suite in Section IV. We vary one parameter once and other parameters are fixed.

1) Sensitivity to N: To test the effect of the population size N, EDA$_{mvn}$ with $N = 100, 200, 300, 500, 600, 1000$ are run on the test instances, while the number of bins $W = 2000$. Fig. 5 shows that with the increment of the population size, the profits increase first from $N = 100$–600 in most curves, while decreasing when $N = 1000$. Fig. 5 illustrates that EDA$_{mvn}$ can find more accurate solutions and get the best solution when $N = 600$, but the solution quality in some of the

TABLE VII
COMPARISON RESULTS ON THE SYNTHETIC BENCHMARK SUITE WITH DIFFERENT DEGREE OF CFS. THE RESULTS SHOWN BY THE NORMALIZED VALUES. EACH ROW SHOWS THE RESULTS OF DIFFERENT ALGORITHMS UNDER A CERTAIN FACTOR LEVEL. THE BEST RESULTS ARE SHOWN IN BOLD TYPE

	CF	EDA_{mvn}	GA-PSO	ACO_{mv}	DE_{mv}	AEDA	Lingo
k_1^3	Weak	**0.889**	0.482	0.112	0.868	0.633	0.264
k_2^3	Strong	0.846	0.639	0.040	**0.902**	0.484	0.331

TABLE VIII
COMPARISON RESULTS ON THE SYNTHETIC BENCHMARK SUITE WITH DIFFERENT PS. THE RESULTS SHOWN BY THE NORMALIZED VALUES. EACH ROW SHOWS THE RESULTS OF DIFFERENT ALGORITHMS UNDER A CERTAIN FACTOR LEVEL. THE BEST RESULTS ARE SHOWN IN BOLD TYPE

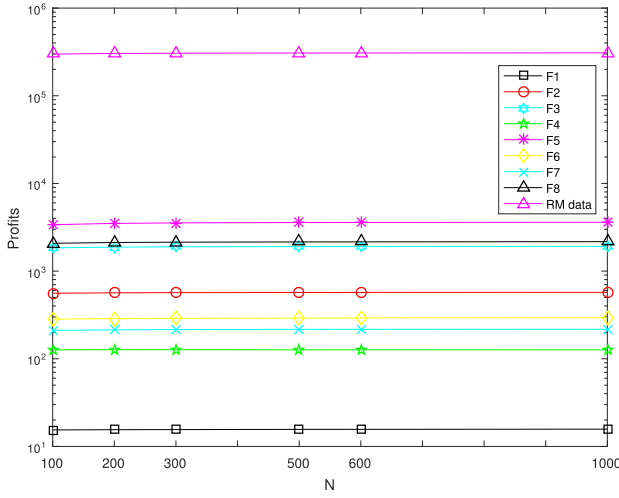|	PS	EDA_{mvn}	GA-PSO	ACO_{mv}	DE_{mv}	AEDA	Lingo
k_1^4	Yes	0.735	0.516	0.051	**0.967**	0.597	0.246
k_2^4	No	**1.000**	0.605	0.100	0.803	0.519	0.349

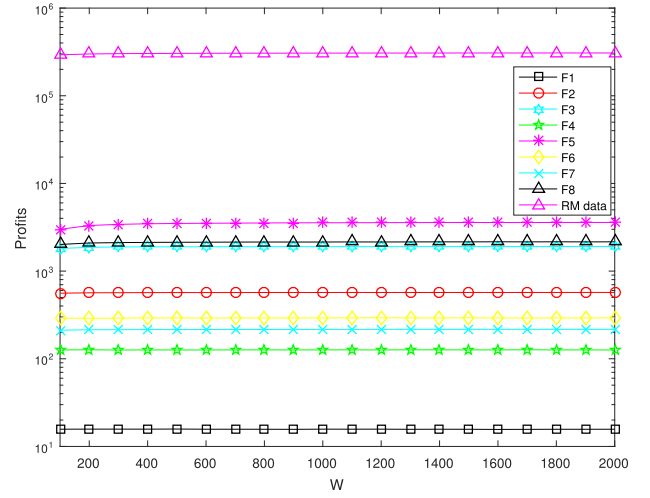Fig. 5. Mean profit for EDA$_{mvn}$ with different population size.

Fig. 6. Mean profit for EDA$_{mvn}$ with different number of bins.

instances becomes worse when $N = 1000$. The reason can be summarized as follows. Since most of our problems are high-dimensional and have a wide search space, a small population size may make it hard to build a high-quality model. On the other hand, if the population size becomes too large, the number of iterations will be reduced under a fixed number of FEs, and the algorithm may not be able to find the optimal solutions.

2) Sensitivity to W: To test the effect of the number of bins W, EDA$_{mvn}$ with $W = 100, 200, \ldots, 1900, 2000$ are tested, and the population size N is set as 600. As the process to build the probability model by the AWH model will become more time-consuming with the increment on W, the maximum number of bins is set as 2000 in our experiment.

The results are shown in Fig. 6. It is obvious that the quality of the optimal solution increases according to the growth of W. In the AWH model, since the bins in the promising area is shrinking as the search goes, the bin that contain the optimal solution is also becoming narrow. Therefore, it is very promising to find high-quality solutions. Although a smaller W can make the bin that contains the optimal solution have a higher probability to be selected, the larger width of the bin will make it more difficult to sample the optimal solution in

this bin. A larger W can partition the search space into lots of narrower bins. In this case, although the number of bins will increase, many of these bins are empty and cannot be selected. Furthermore, a narrower width can increase the probability of sampling the optimal solution in the later random sampling process and it is very helpful to find high precise solutions.

VI. CONCLUSION

Since the newsvendor model has a significant practical application in the field of production, service and management, various variants of the newsvendor model have been proposed. However, previous variants mainly focus on the order quantity but ignore the effect of the real demand on selling price which cannot satisfy the real-world investment market. This article takes into account the price as decision variable and propose a new newsvendor model. Then, the newsvendor problem can be formulated as a constrained mixed-variable nonlinear programming (MINLP) problem.

In order to address the new newsvendor problem, this article designed a mixed-variable EDA, i.e., EDA$_{mvn}$. First, in order to promote the efficiency of initialization, a constraints handling strategy is employed to generate feasible solutions. Second, an

EDA based on histogram models, the AWH model for the continuous variables and LBH model for the discrete variables, is used as the optimizer to find the optimal solution. Besides that, a constraint-based selection method is employed to find high-quality solutions. In order to evaluate the performance of EDA$_{mvn}$, a synthetic benchmark suite is developed by the orthogonal experiment design method and a real-world benchmark is used to further evaluate the performance of EDA$_{mvn}$ in real applications.

The experimental results show that, the proposed EDA$_{mvn}$ algorithm is effective to obtain more profit on the new mixed-variable newsvendor model than the classical model. Furthermore, EDA$_{mvn}$ has been compared with the commercial solver (i.e., Lingo) and other state-of-the-art MINLP algorithms, and also show a competitive performance.

Although there are many researches which try to solve the mixed-variable optimization problem using EAs, little work has been reported on the research of newsvendor problem, especially for MINLP newsvendor problems. This article proposes an EDA-based method for newsvendor problem, which offers a new perspective and an effective way to deal with MINLP problems. First, due to the histogram model-based framework does not require any prior information about distribution of solutions, and the adaptive strategy employed in AWH and LBH models can help to balance the exploration ability and the exploitation ability well, it is worthwhile exploring the proposed EDA$_{mvn}$ to solve other MINLP problems in future studies. Second, the mixed-variable newsvendor model proposed in this article mainly incorporates the profit as the only optimization objective without considering market risk. The trading risk should be taken into account and the multiobjective newsvendor model should be investigated in the future.

REFERENCES

[1] S. C. Chiam, K. C. Tan, and A. A. Mamun, "Investigating technical trading strategy via an multi-objective evolutionary platform," *Expert Syst. Appl.*, vol. 36, no. 7, pp. 10408–10423, 2009.

[2] V. Deshpande and M. Arikan, "The impact of airline flight schedules on flight delays," *Manuf. Service Oper. Manag.*, vol. 14, no. 3, pp. 423–440, 2012.

[3] M. Densing, "Dispatch planning using newsvendor dual problems and occupation times: Application to hydropower," *Eur. J. Oper. Res.*, vol. 228, no. 2, pp. 321–330, 2013.

[4] F. Zaman, S. M. Elsayed, T. Ray, and R. A. Sarkerr, "Evolutionary algorithms for finding Nash equilibria in electricity markets," *IEEE Trans. Evol. Comput.*, vol. 22, no. 4, pp. 536–549, Aug. 2018.

[5] M. A. Pearson, "The incorporation of target performance measures and constrained optimisation in the newsboy problem," *J. Oper. Res. Soc.*, vol. 51, no. 6, pp. 744–754, 2000.

[6] W. Jammernegg and P. Kischka, "Risk-averse and risk-taking newsvendors: A conditional expected value approach," *Rev. Manag. Sci.*, vol. 1, no. 1, pp. 93–110, 2007.

[7] B. Zhang, "Multi-tier binary solution method for multi-product newsvendor problem with multiple constraints," *Eur. J. Oper. Res.*, vol. 218, no. 2, pp. 426–434, 2012.

[8] L. L. Abdel-Malek and N. Areeratchakul, "A quadratic programming approach to the multi-product newsvendor problem with side constraints," *Eur. J. Oper. Res.*, vol. 176, no. 3, pp. 1607–1619, 2007.

[9] M. Cherikh, "On the effect of centralisation on expected profits in a multi-location newsboy problem," *J. Oper. Res. Soc.*, vol. 51, no. 6, pp. 755–761, 2000.

[10] M. Ujjwal, S. Vishakha, T. S. Lakshman, and K. T. Manoj, "Multiple criteria risk averse model for multi-product newsvendor problem using conditional value at risk constraints," *Inf. Sci.*, vol. 478, pp. 595–605, Apr. 2019.

[11] W. Mei, L. Du, B. Niu, J. Wang, and J. Feng, "The effects of an undisclosed regular price and a positive leadtime in a presale mechanism," *Eur. J. Oper. Res.*, vol. 250, no. 3, pp. 1013–1025, 2016.

[12] E. M. Gerber and J. Hui, "Crowdfunding: Motivations and deterrents for participation," *ACM Trans. Comput.–Human Interact.*, vol. 20, no. 6, p. 34, 2013.

[13] G. J. Kyparisis and C. Koulamas, "The price-setting newsvendor problem with nonnegative linear additive demand," *Eur. J. Oper. Res.*, vol. 269, no. 2, pp. 695–698, 2018.

[14] L. Xin, H. Long, J. Bewli, J. Bowman, H. Feng, and Z. Qin, *On the Performance of Tailored Base-Surge Policies: Theory and Application at Walmart.com.* Rochester, NY, USA: Soc. Sci. Electron., 2017.

[15] C. Erbas, S. Cerav-Erbas, and A. D. Pimentel, "Multiobjective optimization and evolutionary algorithms for the application mapping problem in multiprocessor system-on-chip design," *IEEE Trans. Evol. Comput.*, vol. 10, no. 3, pp. 358–374, Jun. 2006.

[16] T. Tometzki and S. Engell, "Systematic initialization techniques for hybrid evolutionary algorithms for solving two-stage stochastic mixed-integer programs," *IEEE Trans. Evol. Comput.*, vol. 15, no. 2, pp. 196–214, Apr. 2011.

[17] Q. N. Huynh, S. Chand, H. K. Singh, and T. Ray, "Genetic programming with mixed integer linear programming based library search," *IEEE Trans. Evol. Comput.*, vol. 22, no. 5, pp. 733–747, Oct. 2018.

[18] L. L. Abdel-Malek and R. Montanari, "On the multi-product newsboy problem with two constraints," *Comput. Oper. Res.*, vol. 32, no. 8, pp. 2095–2116, 2005.

[19] J. Rubio-Herrero and M. Baykal-Gürsoy, "On the unimodality of the price-setting newsvendor problem with additive demand under risk considerations," *Eur. J. Oper. Res.*, vol. 265, no. 3, pp. 962–974, 2018.

[20] L. Yu, W. Fan, J. Pei, and P. M. Pardalos, "A study on the effect of yield uncertainty in price-setting newsvendor models with additive-multiplicative demand," *Optim. Lett.*, vol. 12, no. 6, pp. 1421–1441, 2017.

[21] M. Xu, Y. Chen, and X. Xu, "The effect of demand uncertainty in a price-setting newsvendor model," *Eur. J. Oper. Res.*, vol. 207, no. 2, pp. 946–957, 2010.

[22] S. Zhang and Y. Xia, "Solving nonlinear optimization problems of real functions in complex variables by complex-valued iterative methods," *IEEE Trans. Cybern.*, vol. 48, no. 1, pp. 277–287, Jan. 2018.

[23] Y. Merzifonluoglu and Y. Feng, "Newsvendor problem with multiple unreliable suppliers," *Int. J. Prod. Res.*, vol. 52, no. 1, pp. 221–242, 2014.

[24] B. Zhang, X. Xu, and Z. Hua, "A binary solution method for the multi-product newsboy problem with budget constraint," *Int. J. Prod. Econ.*, vol. 117, no. 1, pp. 136–141, 2009.

[25] R. Li *et al.*, "Mixed integer evolution strategies for parameter optimization," *Evol. Comput.*, vol. 21, no. 1, pp. 29–64, 2013.

[26] K. L. Sadowski, P. A. N. Bosman, and D. Thierens, "A clustering-based model-building EA for optimization problems with binary and real-valued variables," in *Proc. Genet. Evol. Comput. Conf. (GECCO)*, 2015, pp. 911–918.

[27] K. L. Sadowski, P. A. N. Bosman, and D. Thierens, "Learning and exploiting mixed variable dependencies with a model-based EA," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2016, pp. 4382–4389.

[28] L. J. Li, Z. B. Huang, and F. Liu, "A heuristic particle swarm optimization method for truss structures with discrete variables," *Comput. Struct.*, vol. 87, nos. 7–8, pp. 435–443, 2009.

[29] C. L. Sun, J. C. Zeng, and J. S. Pan, "An improved particle swarm optimization with feasibility-based rules for constrained optimization problems," in *Next-Generation Applied Intelligence. IEA/AIE*, vol. 5579. Heidelberg, Germany: Springer, Jun. 2009, pp. 202–211.

[30] T. Liao, K. Socha, M. A. M. D. Oca, T. Stützle, and M. Dorigo, "Ant colony optimization for mixed-variable optimization problems," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 503–518, Aug. 2014.

[31] L. Sahoo, A. Banerjee, A. K. Bhunia, and S. Chattopadhyay, "An efficient GA-CPSO approach for solving mixed-integer nonlinear programming problem in reliability optimization," *Swarm Evol. Comput.*, vol. 19, pp. 43–51, Dec. 2014.

[32] G.-C. Luh, C.-Y. Lin, and Y.-S. Lin, "A binary particle swarm optimization for continuum structural topology optimization," *Appl. Soft Comput.*, vol. 11, no. 2, pp. 2833–2844, 2011.

[33] S. He, E. Prempain, and Q. H. Wu, "An improved particle swarm optimizer for mechanical design optimization problems," *Eng. Optim.*, vol. 36, no. 5, pp. 585–605, 2004.

[34] A. H. Gandomi, X.-S. Yang, and A. H. Alavi, "Mixed variable structural optimization using Firefly algorithm," *Comput. Struct.*, vol. 89, no. 23, pp. 2325–2336, 2011.

[35] M. Urselmann, S. Barkmann, G. Sand, and S. Engell, "A memetic algorithm for global optimization in chemical process synthesis problems," *IEEE Trans. Evol. Comput.*, vol. 15, no. 5, pp. 659–683, Oct. 2011.

[36] Y. Lin, W. Du, T. Liao, and T. Stützle, "Three L-SHADE based algorithms on mixed-variables optimization problems," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2017, pp. 2274–2281.

[37] K. P. Dahal, G. M. Burt, J. R. McDonald, and A. Moyes, "A case study of scheduling storage tanks using a hybrid genetic algorithm," *IEEE Trans. Evol. Comput.*, vol. 5, no. 3, pp. 283–294, Jun. 2001.

[38] W. Shi, W. N. Chen, Y. Lin, T. Gu, S. Kwong, and J. Zhang, "An adaptive estimation of distribution algorithm for multi-policy insurance investment planning," *IEEE Trans. Evol. Comput.*, vol. 23, no. 1, pp. 1–14, Feb. 2019.

[39] A. A. Taleizadeh, S. T. A. Niaki, and V. Hoseini, "Optimizing the multiproduct, multi-constraint, bi-objective newsboy problem with discount by a hybrid method of goal programming and genetic algorithm," *Eng. Optim.*, vol. 41, no. 5, pp. 437–457, 2009.

[40] M. Keramatpour, S. T. A. Niaki, and S. H. R. Pasandideh, "A bi-objective two-level newsvendor problem with discount policies and budget constraint," *Comput. Ind. Eng.*, vol. 120, pp. 192–205, Jun. 2018.

[41] D. Nan, S.-D. Zhou, and Z.-Q. Sun, "Histogram-based estimation of distribution algorithm: A competent method for continuous optimization," *J. Comput. Sci. Technol.*, vol. 23, no. 1, pp. 35–43, 2008.

[42] X. Jing, Y. P. Yan, and J. Zhang, "*HPBILc*: A histogram-based EDA for continuous optimization," *Appl. Math. Comput.*, vol. 215, no. 3, pp. 973–982, 2009.

[43] Y. Hasegawa and H. Iba, "Latent variable model for estimation of distribution algorithm based on a probabilistic context-free grammar," *IEEE Trans. Evol. Comput.*, vol. 13, no. 4, pp. 858–878, Aug. 2009.

[44] X. Li, S. Mabu, and K. Hirasawa, "A novel graph-based estimation of the distribution algorithm and its extension using reinforcement learning," *IEEE Trans. Evol. Comput.*, vol. 18, no. 1, pp. 98–113, Feb. 2014.

[45] Y. Chen, X. Y. Sun, D. W. Gong, Y. Zhang, J. Choi, and S. Klasky, "Personalized search inspired fast interactive estimation of distribution algorithm and its application," *IEEE Trans. Evol. Comput.*, vol. 21, no. 4, pp. 588–600, Aug. 2017.

[46] J. Ceberio, E. Irurozqui, A. Mendiburu, and J. A. Lozano, "A distance-based ranking model estimation of distribution algorithm for the flow-shop scheduling problem," *IEEE Trans. Evol. Comput.*, vol. 18, no. 2, pp. 286–300, Apr. 2014.

[47] X. Liang, H. Chen, and J. A. Lozano, "A Boltzmann-based estimation of distribution algorithm for a general resource scheduling model," *IEEE Trans. Evol. Comput.*, vol. 19, no. 6, pp. 793–806, Dec. 2015.

[48] A. Zhou, J. Sun, and Q. Zhang, "An estimation of distribution algorithm with cheap and expensive local search methods," *IEEE Trans. Evol. Comput.*, vol. 19, no. 6, pp. 807–822, Dec. 2015.

[49] P. Yang, K. Tang, and X. Yao, "Turning high-dimensional optimization into computationally expensive optimization," *IEEE Trans. Evol. Comput.*, vol. 22, no. 1, pp. 143–156, Feb. 2018.

[50] Z. Li and Q. Zhang, "A simple yet efficient evolution strategy for large scale black-box optimization," *IEEE Trans. Evol. Comput.*, vol. 22, no. 5, pp. 637–646, Oct. 2018.

[51] L. Abdel-Malek, R. Montanari, and L. C. Morales, "Exact, approximate, and generic iterative models for the multi-product newsboy problem with budget constraint," *Int. J. Prod. Econ.*, vol. 91, no. 2, pp. 189–198, 2004.

[52] H. Hongfu, H. Yong, and L. Dong, "Coordination of pricing, inventory, and production reliability decisions in deteriorating product supply chains," *Int. J. Prod. Res.*, vol. 56, no. 18, pp. 6201–6224, 2018.

[53] Y. Lin, Y. Liu, W. N. Chen, and J. Zhang, "A hybrid differential evolution algorithm for mixed-variable optimization problems," *Inf. Sci.*, vol. 466, pp. 170–188, Oct. 2018.
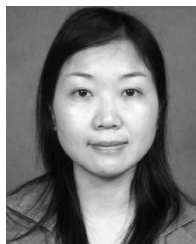
Feng Wang (M'09) received the B.Sc. and Ph.D. degrees in computer science from Wuhan University, Wuhan, China, in 2003 and 2008, respectively.

She is currently an Associate Professor with the School of Computer Science, Wuhan University. Her current research interests include evolutionary computation, intelligent information retrieval, and machine learning.

Dr. Wang serves as a reviewer for several international journals and conferences.
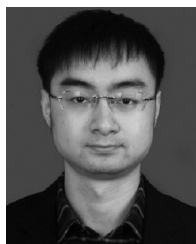
Yixuan Li received the B.Eng. degree in engineering from Wuhan University, Wuhan, China, in 2015, where he is currently pursuing the postgraduate degree.

His current research interests include multiobjective evolutionary computation and evolutionary learning.

Aimin Zhou (S'08–M'10–SM'17) received the B.Sc. and M.Sc. degrees from Wuhan University, Wuhan, China, in 2001 and 2003, respectively, and the Ph.D. degree from the University of Essex, Colchester, U.K., in 2009, all in computer science.

He is currently a Professor with the School of Computer Science and Technology, East China Normal University, Shanghai, China. He has authored over 70 peer-reviewed papers. His current research interests include evolutionary computation and optimization, machine learning, image processing, and their applications.

Prof. Zhou is an Associate Editor of *Swarm and Evolutionary Computation*, and an Editorial Board Member of *Complex and Intelligent Systems*.

Ke Tang (M'07–SM'13) received the B.Eng. degree from the Huazhong University of Science and Technology, Wuhan, China, in 2002, and the Ph.D. degree from Nanyang Technological University, Singapore, in 2007.

From 2007 to 2017, he was with the School of Computer Science and Technology, University of Science and Technology of China, Hefei, China, first as an Associate Professor from 2007 to 2011 and later as a Professor from 2011 to 2017. He is currently a Professor with the Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China. He has published over 60 journal papers and over 80 conference papers. He has over 6000 Google Scholar citations with an *H*-index of 36. His current research interests include evolutionary computation, machine learning, and their applications.

Dr. Tang was a recipient of the Royal Society Newton Advanced Fellowship in 2015 and the 2018 IEEE Computational Intelligence Society Outstanding Early Career Award. He is an Associate Editor of the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION and served as an Editorial Board Member for a few other journals.