A Prediction Strategy based on Decision Variable Analysis for Dynamic Multi-objective Optimization

Jinhua Zheng\textsuperscript{a,b}, Yubing Zhou\textsuperscript{a,*}, Juan Zou\textsuperscript{a}, Shengxiang Yang\textsuperscript{c}, Junwei Ou\textsuperscript{a}, Yaru Hu\textsuperscript{a}

\textsuperscript{a}Key Laboratory of Intelligent Computing and Information Processing (Ministry of Education), Xiangtan University, Xiangtan, Hunan, 411105, China
\textsuperscript{b}Hunan Provincial Key Laboratory of Intelligent Information Processing and Application, Hengyang Normal University, Hengyang, 421002, China
\textsuperscript{c}School of Computer Science and Informatics, De Montfort University, Leicester LE1 9BH, U.K.

Abstract

Many multi-objective optimization problems in reality are dynamic, requiring the optimization algorithm to quickly track the moving optima after the environment changes. Therefore, response strategies are often used in dynamic multi-objective algorithms to find Pareto optimal. In this paper, we propose a hybrid prediction strategy based on the classification of decision variables, which consists of three steps. After detecting the environment change, the first step is to analyze the influence of each decision variable on individual convergence and distribution in the new environment. The second step is to adopt different prediction methods for different decision variables. Finally, adaptive selection is applied to the solution set generated in the first and second steps, and solutions with good convergence and diversity are selected to make the initial population more adaptable to the new environment. The prediction strategy can help the solution set converge while maintaining its diversity. The experimental results and performance show that the proposed algorithm is capable of significantly improving the dynamic optimization performance compared with five state-of-the-art evolutionary algorithms.

\textsuperscript{*}Corresponding author

Email addresses: jzheng@xtu.edu.cn (Jinhua Zheng), 872258528@qq.com (Yubing Zhou), zoujuan@xtu.edu.cn (Juan Zou), syang@dmu.ac.uk (Shengxiang Yang), junweiou@163.com (Junwei Ou), huyaru1999@gmail.com (Yaru Hu)
Keywords:
Dynamic multi-objective optimization, Evolutionary algorithms, Decision Variable Analysis, Adaptive Selection, Diversity

1. Introduction

Many real-world multi-objective optimization problems (MOPs) are dynamic, meaning their objectives conflict with each other and/or parameter may change over time. Over the past decades, many multi-objective evolutionary algorithms (MOEAs) have achieved success on MOPs. However, when faced with complex dynamic multi-objective optimization problems (DMOPs) [1], traditional MOEAs [2] [3] [4] have obvious deficiencies. In DMOPs, the Pareto optimal front (POF) and the Pareto optimal set (POS) are not always fixed, and the optimal population that evolves in a traditional MOEA may lose diversity. Especially in the later stages of evolution, the population gradually converges, which means that the population has difficulty adapting to the new environment. In order to effectively track the changing POF and/or POS over time, MOEAs should be able to respond quickly to environmental changes.

Since MOEAs cannot effectively solve DMOPs, many researchers have done a lot of work on DMOPs, such as detecting whether the environment changes [5] [6] [7] [8]. Some dynamic multi-objective optimization algorithms (DMOEAs) have also been proposed to track a moving POF or POS quickly and obtain a POS that is uniformly distributed over time. On the other hand, in recent years, DMOEAs have been extensively applied in many areas, such as scheduling [9] [10], control [11] [12], planning [13] [14] and design [15], for example, Deb et.al [9] combines NSGA-II with response strategy to effectively solve the hydro-thermal power scheduling. In a hydro-thermal power scheduling, the demand of power is unequal for all concerned units in different time, so the change of the problem is dynamic over time and the solutions of problem must be found when there is a change in power demand. Compared with traditional MOEAs, the algorithm with response strategy can respond to environmental changes more quickly and
help the algorithm jump out of local optimum. Most existing DMOEAs are constructed by combining classic MOEAs with efficient dynamic techniques, including predictive-based and diversity-based approaches. These technologies are competitive in addressing DMOPs, but each technology has usually been limited to solving specific types of problems. For example, prediction-based methods typically achieve better performance on DMOPs with predictable characteristics. Diversity-based approaches have advantages in dealing with DMOPs with dynamically changing characteristics that can lead to a severe loss of diversity. Each method has unique advantages for a specific type of problem.

However, if only the existing diversity generation strategies are adopted, we can not effectively guide the population to explore decision space, and the convergence of the algorithm is greatly reduced. If only the prediction method is used, the performance of the algorithm depends on the training results of the prediction model. In order to avoid these defects as much as possible and combine the advantages of the two methods, we propose a prediction strategy based on decision variable analysis (DVA). When an environmental change is detected, the influence of each decision variable on individual convergence and distribution in the new environment is analyzed, and then combined with historical information, different strategies are adopted to re-initialize the different decision variables. Finally, using adaptive selection to select the initial population with better convergence and distribution solutions for the new environment, we hope that the prediction strategy can ensure that the population can adapt to the environment quickly while maintaining good diversity. Experimental results show that DVA is more adaptable to changing environments than the five algorithms. The main contributions of this study are summarized as follows.

- A decision variable analysis method is used to effectively analyze the impact of decision variables on individuals in the new environment. Combining with historical information, the initial population is generated in different ways according to different decision variables, which can avoid the loss of diversity when the prediction strategy promotes population
convergence.

- We propose an angle-based adaptive selection method to select solutions with better convergence and distribution to form the initial population, so as to better balance the convergence and distribution of the population.

- DVA is a response strategy that can be combined with other classical algorithms to effectively handle DMOPs.

The paper is organized as follows. Section II provides a review of related work on DMOPs and motivation. The proposed prediction strategy based on DVA and the general framework for the proposed method is detailed in Section III. Section IV introduces the test instances, compared algorithms, performance metric and parameter settings. The results and discussions are provided in Section V and Section VI. Finally, conclusions are drawn in Section VII.

2. Related Work

In this paper, we consider that minimization problems and DMOPs [1] [16] [17] can be presented as follows:

\[
\begin{align*}
\min F(x, t) &= \{f_1(x, t), f_2(x, t), \ldots, f_m(x, t)\}, \\
\text{s.t.} g(x, t) &\leq 0, h(x, t) = 0,
\end{align*}
\]

where \(x = (x_1, x_2, \ldots, x_n)\) is the \(n\)-dimensional decision variables within the decision space \(\Omega\); \(F = (f_1, f_2, \ldots, f_m)\) is the \(m\)-dimensional objective vector; \(g(x, t) \leq 0\) and \(h(x, t) = 0\) are the inequality and equality constraints. The time variable \(t\) [18] is associated with the generation number of the EA, and can be calculated as follows:

\[
t = \frac{1}{n_t} \lceil \frac{\tau}{\tau_t} \rceil,
\]

where \(\tau\) is the generation number; \(n_t\) is change severity, and \(\tau_t\) is change frequency. There are some definitions of DMOPs as follows:
Definition 1. Pareto Dominance [2]: Assume that \( p \) and \( q \) are any two individuals in the population; \( p \) is said to dominate \( q \), written as \( f(p) \succ f(q) \), if \( f_i(p) \leq f_i(q) \ \forall i \in 1, 2, \ldots, m \) and \( f_j(p) < f_j(q) \ \exists j \in 1, 2, \ldots, m \).

Definition 2. Pareto Optimal Set (POS): \( x \) is the decision vector; \( \Omega \) is the decision space; \( F \) is the objective function. A solution is said to be nondominated if it is not dominated by any other solutions in \( \Omega \). Thus, the POS is the set of all nondominated solutions and can be defined mathematically as follows:

\[
POS := \{ x \in \Omega | \nexists x^* \in \Omega, F(x^*) \prec F(x) \}.
\] (3)

Definition 3. Pareto Optimal Front (POF): \( x \) is the decision vector; \( \Omega \) is the decision space; \( F \) is the objective function. Thus, the POF is the set of all nondominated solutions with respect to the objective space and can be defined mathematically as follows:

\[
POF := \{ y = F(x) | x \in POS \}.
\] (4)

On the other hand, based on the dynamic changes of the POS and POF, Farina et al. [2] classified DMOPs into four different types.

- **Type I**: The POS changes with time but the POF is fixed.
- **Type II**: Both the POS and POF change with time.
- **Type III**: The POS remains fixed, while the POF changes with time.
- **Type IV**: Both the POS and POF remain fixed.

We mainly deal with the first three types of changes in DMOPs, although the Type IV change may also occur in some cases.

In recent years, DMOEAs [19] [20] [21] [22] have been proposed to deal with DMOPs. The majority can be classified as diversity-based approaches, convergence-based approaches or other approaches [23] [18] [24]. Convergence-based approaches focus on enhancing population convergence. Current convergence-based methods include the memory-based strategy [25] [26] [27] [28], the predictive strategy [21] [30] [31] [32] [33] [34]. Memory-based methods respond quickly
to environmental changes by recording past historical information. When POF
or POS changes periodically, it usually achieves the best results. On the other
hand, The predictive strategy responds to environmental changes by combining
other methods, such as strategies in machine learning [35] [22]. However, it
usually requires a training cycle, and the performance of such models may be
poor.

Unlike convergence-based approaches, diversity-based methods focus on main-
taining the population diversity. When an environmental change is detected,
the decision variable space can be better searched. Generally, diversity-based
approaches can be divided into two categories according to the way diversity is
enhanced; these are diversity introduction and diversity preservation. Diversity
introduction method [36] [37] [38] is mainly used to respond to environmental
changes. Specifically, when environmental changes are detected, diversity in-
troduction method is used to generate some solutions to increase the diversity
of the population. The latter method [39] [23] pays attention to the diversity
preservation of the algorithm itself. Such methods rely only on static evolution
capabilities to find a set of optimal solutions, so the convergence ability of the
algorithm is lacking. Among many kinds of DMOEAs, we mainly consider the
following two categories:

A. Diversity Introduction Methods

The diversity introduction method mainly considers the loss of potential di-
versity in a dynamic environment, especially at a later stage of an environmental
moment, when the population almost converges. If an environmental change is
detected, the population loses its ability to explore the entire decision space due
to convergence to the POS, so the population cannot quickly converge to the
new POS. In order to solve this problem, many DMOEAs have been proposed in
recent years. For example, Deb et al. extended NSGA-II [2] into two algorithm-
s (DNSGA-II-A and DNSGA-II-B) [9]. When a change in the environment is
detected, DNSGA-II-A will randomly re-initialize 20% of the individuals, and
another algorithm DNSGA-II-B will randomly mutate 20% of the individuals,
both of which can increase the diversity of the population to some extent. Harri-
son et al. [40] proposed dynamic vector evaluated particle swarm optimization. Based on this, Helbig et al. [41] proposed a heterogeneous dynamic vector evaluated particle swarm optimisation (HDVEPSO) algorithm. When the objective function changes, HDVEPSO randomly re-initializes 30 percent of the particles to avoid falling into local optimum. Goh et al. [39] introduced a competitive-cooperative coevolutionary algorithm (dCOEA) where random individuals are generated to enhance diversity of the population if the environment changes. In conclusion, these diversity introduction methods randomly generate or mutate individuals to prevent the population from falling into local optimum after detecting changes in the environment. However, most of these methods increase population diversity by randomly reinitializing or mutating only some of the individuals, and cannot effectively explore the new decision space. Additionally, the historical information cannot be used effectively, so the algorithm does not converge fast.

B. Prediction-based Methods

The prediction strategy uses historical information to generate an initial population to help the population adapt quickly to the environment. For different DMOPs, a suitable predictive model is critical. A good predictive model can provide a guiding direction for the evolution of the population towards the POF. Hatzakis and Wallace [42] proposed a feed-forward prediction strategy (FPS) that predicts only the boundary points of the population, which makes it difficult to deal with DMOPs with nonlinear correlations between decision variables. Zhou and Jin et al. [43] proposed a population prediction strategy (PPS). Both FPS and PPS use the autoregressive model to predict population. In addition, due to the lack of historical information in the early stage, PPS convergence is very slow. Ruan et al. [44] applied gradual search to predict the idea position of the individual in the new environment. Jiang et al. [45] suggested a steady-state and generational evolutionary algorithm (SGEA), which guides the search of solutions by a moving direction from the centroid of the non-dominated solution set to the centroid of the whole population. Recently, Cao et al. [34] presents a novel prediction model combined with a multiobjective
evolutionary algorithm based on decomposition to solve dynamic multiobjective optimization problems. In short, these prediction strategies can enhance the convergence speed of the population in the new environment to a certain extent. However, there may be problems in predicting so errors may occur occasionally and some prediction strategies require additional training processes, which consumes a lot of computing resources.

From this discussion, we can see that the two methods have their own advantages in resolving different aspects of DMOPs. On the premise of introducing the method of DVA, in this study, we hybridize the diversity introduction methods and the prediction-based methods to improve the quality of solutions. The proposed algorithm is described next.

3. Proposed Algorithm

In this section, we propose a new algorithm for solving DMOPs, which aims to generate a population close to the true $POF$ while maintaining the ability of the population to explore the decision space. The main idea is to analyze the impact of each decision variables on individuals at the current environment and to adopt different strategies to optimize them. To achieve this target, we use a simple decision variable analysis method when the environment changes, and then adopt different generation methods for different decision variables to generate the initial population.

3.1. Decision variable analysis method

The DVA method analyzes the influence of each decision variables on individuals and then performs simple classification processing when the change is detected. The purpose is to find distribution-related decision variables, and other decision variables are considered to be related to convergence. Therefore, we randomly generate $n$ representative solutions representing $n$-dimensional decision variables, and then disturb the decision variables represented by each representative solution. The values of other decision variables remain unchanged.
Each representative solution generates \( nPer \) perturbation solutions through perturbation. Finally, the representative solution and its perturbation solution are ranked by nondominated sort [2]. We judge the influence of the dimension decision variables on individuals in the current environment by the dominance relationship between the representative solution and the perturbation solution.

Fig. 1 presents an example to illustrate the main idea of the DVA, where a bi-objective minimization problem with five decision variables \( x_1, x_2, x_3, x_4 \) and \( x_5 \) is considered. A number of \( n \) representative solutions (five in this example) are first randomly generated. Then, a number of \( nPer \) (six in this example) perturbations are performed on each of the five variables of the representative solutions. Fig 1(a) illustrates the objective value of the representative solutions and the solutions generated by the perturbations. A nondominated sort is performed for each representative solution and the perturbation solution it generates. As shown in Fig 1(a), there are three situations:

- The relationship between the representative solution and the disturbance solution is nondominated (the gray and black points in Fig 1(a)), so the decision variables representing the dimension can only affect the distribution of individuals in the current environment, but cannot affect the convergence of individuals.
• The relationship between the representative solution and the perturbation solution is dominated (the blue and white points in Fig. 1(a)), so the dimensional decision variable can only affect the individual convergence under the current environment, but cannot affect the individual distribution.

• Both dominated and nondominated relations exist between the representative solution and the perturbation solution (the orange points in Fig. 1(a)). The representative decision variable can affect both the distribution and the convergence of the individual.

For these three cases, the decision variables in the first and second cases can affect the individual distribution, so the dimensional decision variables are classified as distributed correlation (DV), and the other decision variables are classified as the set of other decision variables (OV, which can only affect the individual convergence), as shown in figure 1(b). The specific algorithm flow is described in detail in algorithm 1.

3.2. The Prediction Strategy Based On DVA

The prediction strategy in this section is to find the solution set as close as possible to the true POS when a change is detected. Different from other prediction strategies, in this paper, different prediction methods are used according to different decision variables to generate a set with sufficient capacity to explore the decision space and move closer to the true POS. The proposed algorithm determines the strategy adopted by the per-dimensional decision variables in the current environment based on the information obtained from the DVA of the previous moment and the current time environment. If the $i_{th}$ decision variable belongs to $OV_t$ at the previous moment and the current moment belongs to $DV_{t+1}$, or the previous moment belongs to $DV_t$ and the current moment belongs to $OV_{t+1}$, then the historical information on the $i_{th}$ is useless, and the strategy of random initialization is adopted on this dimension. In other cases, different prediction mechanisms are adopted according to the classification of
Algorithm 1 Decision Variable Analysis Method

Input:
\[ n \] (the number of decision variable at \( t + 1 \) environment time)

Output:
\[ DV \] (Diversity-related), \[ OV \], \[ C \] (a set of perturbations)

1: for \( i = 1 \) to \( n \) do
2: \( p \) (the representative solution) randomly generate, representing the \( i \)-dimensional decision variable.
3: \( C_{\text{tmp}} \) (Temporary sets for storing representative solutions and perturbation solutions) \( \leftarrow \{p\} \cup \) Generated \( n \text{Per} \) solutions by random perturbation of \( P \) in the \( i_{th} \) dimension
4: Sort \( C_{\text{tmp}} \) using nondominated sort
5: if \( C_{\text{tmp}} \) contains non-dominated relationships then
6: \( DV \leftarrow DV \cup \{i\} \);
7: else
8: \( OV \leftarrow OV \cup \{i\} \);
9: end if
10: \( C \leftarrow C \cup C_{\text{tmp}} \) and empty \( C_{\text{tmp}} \)
11: end for

the current time. The details of the prediction method are given as follows. For the decision variables belonging to \( DV_{t+1} \), a hybrid strategy combining Latin hypercube sampling \cite{46} and random sampling is adopted to enhance the exploration ability of the population in the new environment. In Fig.\textsuperscript{2} an example is presented to illustrate the idea of the hybrid strategy.

This strategy mainly uses the minimum point \((min_i)\) and the maximum point \((max_i)\) of the population. Defined as follows:

\[
\begin{cases}
min_i = (min_1^i, min_2^i, ..., min_n^i) \\
max_i = (max_1^i, max_2^i, ..., max_n^i).
\end{cases}
\]

where \( min_i \) and \( max_i \) are the minimum and maximum values of \( i_{th} \) decision
variable at time $t$. Then the moving direction of the minimum and maximum points in $i_{th}$ dimension is defined as

$$\begin{align*}
\min D^i_t &= \min^i_t - \min^i_{t-1} \\
\max D^i_t &= \max^i_t - \max^i_{t-1}.
\end{align*}$$

(6)

Then, the minimum and maximum points at time $t+1$ can be calculated from the minimum and maximum points at time $t$ and the direction of movement.

$$\begin{align*}
\min^i_{t+1} &= \min^i_t + \min D^i_t \\
\max^i_{t+1} &= \max^i_t + \max D^i_t.
\end{align*}$$

(7)

In the space composed of the minimum point and the maximum point at time $t + 1$, the Latin hypercube sampling method is adopted to generate
Figure 3: The results of selecting 10 points from 0 to 1 by random sampling and Latin hypercube sampling.

$N_2$ solutions, and then $N_3$ solutions are randomly generated in the space. The process of Latin hypercube sampling and random sampling is described in Fig.3. Different from random sampling, the Latin hypercube sampling guarantees that there are sample points on every aliquot of each dimension. The combination of the two methods can enhance the exploring ability of the initial population in the predicted decision space. If the i-th decision variable belongs to $OV_{t+1}$ in the current environment, that is, the dimension decision variable can only affect the convergence of the individual at the current time, then in this dimension we use the central point prediction method to accelerate the convergence speed of the population in the new environment. Suppose $C_t$ is the center point of $POS_t$, and $POS_t$ is the $POS$ of the population that is finally optimized at time $t$, then $C_t$ can be calculated by the following formula:

$$C_t = \frac{1}{|POS_t|} \sum_{x \in POS_t} x_t.$$  \hspace{1cm} (8)

$|POS_t|$ is the number of solutions in $POS$ at time $t$; $x_t = x_{1t}, x_{2t}, ..., x_{nt}$ is the solution at time $t$. Therefore, the moving direction $D_t$ of the center point at
time $t$ is defined as follows:

$$D_t = C_t - C_{t-1}.$$  

(9)

Then the value at time $t + 1$ on the dimension can be generated from the center point and the prediction direction at time $t$ according to the following formula.

$$x_{t+1} = x_t + D_t.$$  

(10)

$x_t$ is the final solution obtained before the environmental change, and $x_{t+1}$ is the predicted solution. The flow of predicting the generation of solution sets is described in algorithm 2 in detail. First, $N_1$ solutions are generated by the Latin hypercube sampling method combined with the centre point prediction method (Algorithm 2 Lines 1-14), and then $N_2$ solutions are generated by the random sampling method combined with the centre point prediction method (Algorithm 2 Lines 15-28). The predicted population consists of both solutions.

3.3. Adaptive Selection Based On The Angle

The population has good diversity in the decision space does not mean it also has good diversity in the objective space, for some problems (e.g., the deflection problem), the set with good distribution in the decision space will lead to the deterioration of the population in the objective space. Therefore, when the environment changes, in order to allow the initialized population to adapt to the new environment quickly, and to promote population convergence in the objective space while strengthening its distribution, we propose an angle-based adaptive selection method. After the analysis of decision variables and the prediction based on the analysis, the populations $C$ and $Q$ generated by algorithms 1 and 2 are merged into population $G$. Then the $\cos(\theta)$ between each individual and other individuals in population $G$ and the crowding of each individual are calculated (Algorithm 3 Lines 1-3). We use nondominated sort to accelerate the convergence of the population and select individuals from the
Algorithm 2 The Prediction Strategy Based On DVA

Input:

\[ DV_t, DV_{t+1}, OV_t, OV_{t+1} \]

Output:

\( Q( \text{ a set of individuals generated by prediction } ) \)

for \( i = 1 \) to \( N_2 \) do

\( \text{for } j = 1 \) to \( n \) do

\( \text{if } ((j \in DV_t \& \& j \in OV_{t+1}) || (j \in OV_t \& \& j \in DV_{t+1})) \text{ then} \)

\( x^i_j \text{ randomly generated} \)

else

\( \text{if } (j \in DV_{t+1}) \text{ then} \)

\( x^i_j \text{ Generated by Latin hypercube sampling} \)

else

\( x^i_j \text{ Calculated by formula }[9] \)

end if

end if

\( Q \leftarrow Q \cup \{i\} \)

end for

for \( i = 1 \) to \( N_3 \) do

\( \text{for } j = 1 \) to \( n \) do

\( \text{if } ((j \in DV_t \& \& j \in OV_{t+1}) || (j \in OV_t \& \& j \in DV_{t+1})) \text{ then} \)

\( x^i_j \text{ randomly generated} \)

else

\( \text{if } (j \in DV_{t+1}) \text{ then} \)

\( x^i_j \text{ randomly generated} \)

else

\( x^i_j \text{ Calculated by formula }[9] \)

end if

end if

\( Q \leftarrow Q \cup \{i\} \)

end for

end for
first layer to join the initial population in turn. At the same time, we enhance the distribution of the population through the angle between individuals. If the angle between individuals and the selected individuals is less than the threshold \( \sigma \), then these individuals will be ignored in the selection according to the layers. If all the layers are selected and the number of initial population is still less than \( N \), the individuals with small crowding will be selected to join the initial population. Algorithm 3 describes the process of selecting the solutions.

As shown in fig. 4, assuming a population size of 5, first, the \( \cos(\theta) \) between each individual and the others is calculated according to formula 11 and the sum of the first \( k \) maximum values of \( \cos(\theta) \) is taken as the individual’s crowding. As a simple analysis, if there are more individuals in a certain direction, the greater
Algorithm 3 Adaptive Selection Based On Angle

Input:

$C$ (Solution Set Generated by Algorithm 1), $Q$ (Solution Set Generated by Algorithm 2)

Output:

$P$ (the initial population at environment time $t+1$)

$G \leftarrow C \cup Q$

Calculate the $\cos(\theta)$ between each individual and others in the population $G$ according to Equation 3:

3: Calculate the Crowding: the sum of the top $k$ largest $\cos$ values of each individual in $G$ as the individual’s crowding

Sort $G$ using nondominated sort and Get the number of layers $K$

for $i = 1$ to $K$ do

6: Select individual $x \in G_i$ and $P \leftarrow P \cup \{x\}$

   for $j = 1$ to $|G|$ do
       if $(\cos(x, x_j) > \delta)$ then

9: $G \leftarrow G \setminus \{x_j\}$

   $\quad P_{tmp} \leftarrow P_{tmp} \cup x_j$

   end if

12: end for

end for

while $|P| < N$ do

15: Choose the individual $x$ with the minimum crowding from $P_{tmp}$;

   $P \leftarrow P \cup \{x\}$

   $\quad P_{tmp} \leftarrow P_{tmp} \setminus \{x\}$

18: end while
the first $k$ cos values between individuals in this region and other individuals, that is, the greater the crowding of individuals in that direction; then sort all the individuals using nondominated sort. There are three layers in Fig.4. The first layer has blue points; the second layer has green points, and the last layer has orange points. From the first layer of individuals, individual A is selected to join the population. Since the $\cos(\theta)$ between the individuals E, F and A is less than the threshold $\delta$, E and F are ignored. Similarly, individual B is selected and individual G is ignored. Individual C is selected and individual H is ignored. Individual D is selected and individual I is ignored. The individuals are added to the population according to the number of layers in turn. When all the layers have been selected, if the number of individuals in the population is still less than $N$, the ignored individuals will be selected, and the individual J is finally selected.

$$
\cos(\theta) = \frac{\sum_{m=1}^{M} (f_m(x_i) - z_m^s)(f_m(x_j) - z_m^s)}{\sqrt{\sum_{m=1}^{M} (f_m(x_i) - z_m^s)^2} \sqrt{\sum_{m=1}^{M} (f_m(x_j) - z_m^s)^2}}.
$$

(11)

where $z_m^s = (z_1^s, z_2^s, ..., z_M^s)$ is the ideal point, and $z_m^s = \min_{x_i \in Q} f_m(x_i)$, $x_i$ and $x_j$ are two individuals in the population.

3.4. A General Framework

The procedure of the whole algorithm is illustrated in Algorithm 4 in detail. When the environment changes and the response strategy is not used, the impact of the decision variables on the individual is analyzed in the new environment, and then the different strategies in Algorithm 2 are used to target different decision variables based on the analysis results. Finally, the solution sets generated by Algorithms 1 and 2 are adaptively selected to obtain the initial population in the new environment. For optimizing stationary MOP, we adopt the dynamic evolutionary environment model (DEE) [20] because it can better guarantee the convergence and distribution of the algorithm when the target
dimension is not large. Certainly, as a response strategy, it can combine with most classical algorithms to deal with DMOPs and it is mentioned later in the discussion section.

**Algorithm 4 Framework of DVA**

**Input:**
- $N$ (population size)

**Output:**
- $P$ (population)

Initial population $P = x_1, x_2, \cdots, x_N$

**while** stopping criterion not met **do**

**if** change detected and not responded **then**

4: $(DV_{t+1}, OV_{t+1}, C) = \text{DVA}(n)$ (**Algorithm 1**)

Q = the Prediction Strategy Based On DVA (**Algorithm 2**)

$P = \text{ASA}(C, Q)$ (**Algorithm 3**)

end if

8: Optimize population with DEE

end while

3.5. **Computational complexity of the compared algorithms and DVA**

The optimization algorithm consumes the most computational resources of the compared algorithms and DVA. The computational complexity of each optimization algorithm and DVA are analyzed as follows:

1. DNSGA-II: From the original paper of DNSGA-II [9], the optimization algorithm is NSGA-II [2] and the computational resource is spent on nondominated sorting $O(M(2N)^2)$, crowding-distance assignment $O(M(2N)\log(2N))$ and sorting $O(2N\log(2N))$. The overall computational complexity is $O(MN^2)$; $M$ is the number of objectives and $N$ is the population size.

2. PPS: PPS [43] chooses RM-MEDA [17] as the MOEA optimizer. In RM-MEDA, the computational complexity of RM-MEDA includes modeling, reproduction and the selection operator. The modeling cost is $O(nN)$ where
\(n\) is the number of the decision space. The reproduction spends \(O(nK)\) and \(K\) is the number of clusters. The selection operation is the same as in NSGA-II \([2]\). Therefore, the overall computational complexity is \(O(MN^2)\).

(3) DMS: The prediction in DMS requires \(O(N)\) computational complexity, and the gradual search strategy and the random diversity maintenance strategy all take \(O(N)\) computational complexity. Then computational complexity of the overall framework of DMS is mainly spent on the nondominated sort. Therefore, the overall computational complexity of DMS for one response mechanism is \(O(MN^2)\).

(4) SGEA: SGEA \([45]\) is introduced in section \([4.2]\) and it consumes during steady-state evolution and environmental selection. The whole steady-state evolution part takes \(O(MN^2)\) computations and the environmental selection procedure spends \(O(MN^2)\) computations. Therefore, the overall computational complexity is \(O(MN^2)\).

(5) DVA: For the overall framework of each generation, the main computational resource in DVA is consumed by adaptive selection and the prediction strategy. The DVA method also needs computational resources when an environmental change is detected. It requires a total of \(O(nM)\) computations, and the prediction strategy requires \(O(nN)\) computations. The adaptive selection (line 6 of Algorithm \([4]\) requests \(O(M(N_1 + N)^2)\), where \(M\) is the number of objectives, \(N_1\) is the population size of \(C\) in algorithm \([3]\) and \(N\) is the population size. Therefore, the overall computational complexity of DVA in each generation is \(O(M(N_1 + N)^2)\).

4. Experimental Design

4.1. Test instances

In this section, 16 dynamic frequently used multi-objective test problems are adopted to examine the performance of the algorithm. The test problems include the FDA test suite(FDA1-FDA4), DMOP test suite(DMOP1-DMOP3) and JY test suite(JY1-JY9). The FDA test suite, which was proposed by Farina et al \([1]\),
is the basic test problem in DMOPs. The DMOP test suite was extended from the FDA suite by Goh et al. [39]. The two test suites have linearly correlated decision variables and both are the most commonly used test suite in dynamic multi-objectives and are widely used to test the performance of DMOEAs [43] [44] [45]. The JY test suite is the new DMOPs test suite proposed by Jiang et al [17]. More and more researchers pay attention to JY test suite and apply it to test the performance of DMOEA [48] [49]. It not only has linear correlation between decision variables, but also has nonlinear correlation on some issues. The JY test suite also introduces some complex features that are useful for examining the performance of the algorithm.

4.2. Compared Algorithms

We compared the performance of the proposed prediction strategy with five popular DMOEAs: the random Initialization Strategy(RIS), the dynamic non-dominated sorting genetic algorithm II (DNSGA-II) [9], the population prediction strategy (PPS) [43], the effect of diversity maintenance on prediction(DMS) [44] and a steady-state and generational evolutionary algorithm (SGEA) [45].

(1) DNSGA-II: NSGA-II [2] is a classic algorithm for multi-objective optimization. To apply it to DMOPs, Deb et al. modified the NSGA-II algorithm to adapt to different environments. When it is detected that the environment changes, some individuals are randomly initialized, and some individuals are replaced with mutated solutions to adapt to the new environment.

(2) PPS: PPS divides the population into two parts: the center point and manifold. When the environment changes, the univariate autoregression(AR) model is used to predict the position of the next population center point. Similarly, the next manifold is generated from the previous manifold. The initial population at the next environment time is obtained from the predicted center point plus the manifold.

(2) DMS: DMS consists of three parts, exploration based on prediction, a gradual search strategy and a random diversity maintenance strategy. The first part is mainly to accelerate the convergence of the population, while the second
and third parts focus on maintaining the distribution of the population.

(4) SGEA: SGEA can take advantage of the fast and steady tracking ability of steady-state algorithms and the good diversity preservation of generational algorithms for solving DMOPs. Some information from the previous environment and new environment are used for reacting to environmental changes when a change is detected.

4.3. Performance Metrics

In this section, performance metrics, which can evaluate convergence, distribution and diversity of the obtained solution set, are introduced.

1) Generational Distance (GD): Van Veldhuizen [39] [27] presented the GD metric, which measures the population’s convergence. The GD indicator is defined as follows:

$$GD(POF_t, P_t) = \frac{\sum_{v \in P_t} d(POF_t, v)}{|P_t|},$$  \hspace{1cm} (12)

where \(d(POF_t, v) = \min_{u \in POF_t} \sqrt{\sum_{j=1}^{m} (f^v_j - f^u_j)^2}\) is the minimum Euclidian distance between \(v\) and the point in \(POF_t\). \(POF_t\) is a set of uniformly distributed Pareto optimal points in the \(POF\) at time \(t\); \(P_t\) is the solution obtained by the algorithms.

2) Inverted Generational Distance (IGD): IGD [43] [44] is a metric which assesses the convergence and diversity of the obtained solution set. The IGD is calculated as follows:

$$IGD(POF_t, P_t) = \frac{\sum_{v \in POF_t} d(v, P_t)}{|POF_t|},$$  \hspace{1cm} (13)

where \(d(v, P_t) = \min_{u \in P_t} \sqrt{\sum_{j=1}^{m} (f^v_j - f^u_j)^2}\) is the minimum Euclidian distance between \(v\) and the point in \(P_t\). \(POF_t\) is a set of uniformly distributed Pareto optimal points in the \(POF\) at time \(t\); \(P_t\) is the solution obtained by the algorithms. The IGD [43] performance metric is a comprehensive index and is developed to measure the convergence and diversity of the algorithm’s obtained solutions.
3) Schott’s spacing metric (SP): This kind of indicator was developed by Schott [50] to investigate the distribution of the found Pareto front. The SP is calculated as follows:

\[
SP = \sqrt{\frac{1}{|P_t| - 1} \sum_{i=1}^{|P_t|} (D_i - \bar{D})^2},
\]

where \(D_i\) is the Euclidean distance between the \(i\)th member in \(P_t\), and its nearest member in \(P_t\), and \(\bar{D}\) is the average value of \(D_i\). SP measures how evenly the solutions in \(|P_t|\) are distributed.

4.4. Parameter Settings

In order to fairly compare algorithms, most parameters of the compared algorithms were set to the values used in their original papers. The experimental parameters were set as follows. The severity of change was fixed to 10 and the frequency of change was set to 20, 25 and 30; The population size was \(N = 100\) and the dimensions of the test problem’s decision space were \(n = 10\); In general, 5% of the population was randomly selected and re-evaluated to detect environmental changes. Each algorithm ran independently 20 times on all problems, and there were 50 environmental changes. The crossover probability was \(p_c = 0.8\) and the mutation probability was \(p_m = 1/n\) for all algorithms.

- Parameters in DVA: \(N_2 = 0.8N\) and \(N_3 = 0.2N\) in the prediction strategy and \(\delta = 0.9998\) in the adaptive selection strategy.

5. Experimental Results and Analysis

In order to analyze the ability of the algorithm to solve DMOPs, the average IGD, GD and SP results over a series of time windows were obtained and their standard deviation values are presented in Tables 1, 2 and 3 respectively, where the best values obtained by one of five algorithms are highlighted in bold face. Additionally, the Wilcoxon rank-sum test [51] was carried out to indicate significance between different results at the 0.05 significance level.
5.1. Results on FDA and dMOP problems

It can be obtained from Table 1 that DVA achieved the best results on most issues of FDA and dMOP. The IGD metric mainly depends on the closeness, distribution, and coverage of an approximation to the true POF. Smaller IGD values represent better performance of the algorithm. We can use IGD together with SP and GD to deeply and extensively reveal the algorithm’s performance on the test instances. The IGD of DVA was the best in most of the test problems except for FDA4 and DMOP1. This indicates that DVA had a better distribution and convergence than the other methods. For FDA4, the IGD of DVA was the best when \( \tau_t = 20 \) and 25, but when \( \tau_t \) increased to 30, SGEA was better than DVA. It shows that the distribution and convergence of DVA were weaker than SGEA in dynamic changes. As for DMOP1, the IGD of DVA was second only to SGEA, and with the increase of t, DVA was better than SGEA. The conclusion can be drawn that DVA performs moderately on problems like FDA4.

As shown in Table 2, DVA had the smallest values of GD on the FDA1, FDA4, dMOP2 and dMOP3. The smaller values of GD imply that the algorithm had better convergence than the other algorithms. For most problems, DVA significantly performed better than RIS, DNSGA-II, PPS and DMS. However, when compared with SGEA, DVA performed worse on FDA2, FDA3 and dMOP1. The reason is that the POS of FDA2 and dMOP1 remain fixed. SGEA preserves half of the solutions from the last population, so the values of GD on those test problems are the smallest. As for FDA3, the density of solutions on the POF can be varied.

For the SP metric, as can be seen from Table 3, DVA was better than the other five algorithms on FDA1, DMOP2 and DMOP3. However, on FDA2 and dMOP1, the performance of DMS and PPS was significantly better than SGEA and DVA. This is because the POS of FDA2 and dMOP1 are fixed, and DMS and PPS have more advantages in dealing with these problems. On the three-dimensional problem of FDA4, DMS, DVA and SGEA all show better distribution. As for FDA3, the distribution of DMS is obviously better than
the other algorithms because the gradual search strategy and random diversity maintenance strategy of the DMS can help it maintain better diversity on this problem.

In addition to the indicators represented in the table, we provide an evolution curve of the average IGD value of the test examples in Fig. As can be seen, compared with the other algorithms, DVA can respond to changes more quickly, and had better stability in most test problems, especially in FDA4, so it was able to obtain higher convergence performance. However, on dMOP2 and dMOP3, the stability of DVA was not as good as that of SGEA. The reason may be that SGEA is very capable of tracking environmental changes, whereas DVA is weaker. However, compared with the other algorithms, DVA still had better stability and convergence. Overall, DVA achieved good performance in the FDA and dMOP test suites.

5.2. Results on JY problems

Unlike the FDA and dMOP test suites, JY problems are a new benchmark suite with several complex characteristics including a nonmonotonic and time-varying relationship among decision variables. In addition, the types of some problems change over time in the optimization process. The following observations are drawn from Table. DVA significantly performed best over other approaches on most JY problems except for JY6 and JY7 in terms of the IGD metric. JY6 is a multimodal problem, where not only the number of local optima changes over time, but also the POS is dynamically shifted. Similarly, JY7 is a multimodal problem, and different from JY6, the number of local optima remains fixed, and the overall POF shape can be concave or convex. These tests mainly tests the search ability of the algorithm in the whole decision space. Combined with SP in Table, compared with other algorithms, it is clear that DVA has good distribution on JY6 and JY7. At the same time, from the convergence metric GD in Table, we can see that the convergence of DVA on these two issues is far less than that of SGEA. This also results in DVA being inferior to SGEA in the comprehensive metric IGD. The reason is that DVA cannot
Table 1: Mean and SD of IGD indicator obtained by six algorithms.

<table>
<thead>
<tr>
<th>Prob.</th>
<th>H/D</th>
<th>RIS</th>
<th>DNSGA-II</th>
<th>PPS</th>
<th>DMS</th>
<th>SGEA</th>
<th>DVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDA1</td>
<td>(20, 10)</td>
<td>6.38 (2.52)</td>
<td>1.52 (2.00)</td>
<td>1.78 (2.14)</td>
<td>2.52 (2.87)</td>
<td>2.82 (1.98)</td>
<td>2.70 (1.78)</td>
</tr>
<tr>
<td></td>
<td>(20, 10)</td>
<td>6.38 (2.52)</td>
<td>1.52 (2.00)</td>
<td>1.78 (2.14)</td>
<td>2.52 (2.87)</td>
<td>2.82 (1.98)</td>
<td>2.70 (1.78)</td>
</tr>
<tr>
<td>FDA2</td>
<td>(20, 10)</td>
<td>3.70 (2.97)</td>
<td>3.44 (2.18)</td>
<td>3.14 (2.18)</td>
<td>2.92 (3.50)</td>
<td>2.90 (2.66)</td>
<td>3.18 (2.14)</td>
</tr>
<tr>
<td></td>
<td>(20, 10)</td>
<td>3.70 (2.97)</td>
<td>3.44 (2.18)</td>
<td>3.14 (2.18)</td>
<td>2.92 (3.50)</td>
<td>2.90 (2.66)</td>
<td>3.18 (2.14)</td>
</tr>
<tr>
<td>FDA3</td>
<td>(20, 10)</td>
<td>1.38 (0.78)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
</tr>
<tr>
<td></td>
<td>(20, 10)</td>
<td>1.38 (0.78)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
</tr>
<tr>
<td>FDA4</td>
<td>(20, 10)</td>
<td>1.38 (0.78)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
</tr>
<tr>
<td></td>
<td>(20, 10)</td>
<td>1.38 (0.78)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
<td>1.38 (0.85)</td>
</tr>
</tbody>
</table>

take into account the convergence well while guaranteeing the distribution in the multimode problem.

From Table 2, it can be seen that DVA had better convergence performance on most JY problems. However, on JY4, the performance convergence of DVA was inferior to SGEA. As for JY6 and JY7, the performance of DVA was worse than SGEA. The reason is shown in Table 3. From the distribution indicator of Table 3, we see that the distributions of DVA in JY5 and JY8 test were not as good as those of SGEA and DMS. The reason is similar to that of FDA2 and dMOP1, the POS of JY5 and JY8 are fixed. As mentioned, DVA cannot treat this kind of problem as well as SGEA and DMS, but DVA was still
better than the rest of the algorithms on this problem.

Similarly, in addition to the indicators shown in the table, we provide the evolution curve of the average IGD value of the JY test problem in Figure 6. We can see that, DVA was able to respond to environmental changes fast and stably in most cases. On JY3, JY4 and JY5, SGEA has roughly similar curves, and both DVA and SGEA were superior to the other algorithms. However, on JY6 and JY7, DVA was significantly inferior to SGEA. DVA’s performance was similar to NSGA-II and significantly better than RIS, DMS and PPS. The specific reasons are as analyzed above. For other issues on JY test suites, DVA did greater advantages.

Table 2: Mean and SD of GD indicator obtained by six algorithms.

<table>
<thead>
<tr>
<th>Problem (τₜ, nₜ)</th>
<th>DSSGA-II</th>
<th>PPS</th>
<th>NDSS</th>
<th>DVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDA (25, 10)</td>
<td>3.60e-1</td>
<td>3.00e-1</td>
<td>1.20e-1</td>
<td>1.60e-1</td>
</tr>
<tr>
<td>FDA (30, 10)</td>
<td>4.50e-1</td>
<td>3.80e-1</td>
<td>1.80e-1</td>
<td>2.20e-1</td>
</tr>
<tr>
<td>FDA (35, 10)</td>
<td>5.40e-1</td>
<td>4.80e-1</td>
<td>2.10e-1</td>
<td>2.60e-1</td>
</tr>
<tr>
<td>FDA (40, 10)</td>
<td>6.30e-1</td>
<td>5.80e-1</td>
<td>2.50e-1</td>
<td>3.10e-1</td>
</tr>
<tr>
<td>FDA (45, 10)</td>
<td>7.20e-1</td>
<td>6.80e-1</td>
<td>3.00e-1</td>
<td>3.70e-1</td>
</tr>
</tbody>
</table>

Table 2: Mean and SD of GD indicator obtained by six algorithms.

<table>
<thead>
<tr>
<th>Problem (τₜ, nₜ)</th>
<th>DSSGA-II</th>
<th>PPS</th>
<th>NDSS</th>
<th>DVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDA (25, 10)</td>
<td>3.60e-1</td>
<td>3.00e-1</td>
<td>1.20e-1</td>
<td>1.60e-1</td>
</tr>
<tr>
<td>FDA (30, 10)</td>
<td>4.50e-1</td>
<td>3.80e-1</td>
<td>1.80e-1</td>
<td>2.20e-1</td>
</tr>
<tr>
<td>FDA (35, 10)</td>
<td>5.40e-1</td>
<td>4.80e-1</td>
<td>2.10e-1</td>
<td>2.60e-1</td>
</tr>
<tr>
<td>FDA (40, 10)</td>
<td>6.30e-1</td>
<td>5.80e-1</td>
<td>2.50e-1</td>
<td>3.10e-1</td>
</tr>
<tr>
<td>FDA (45, 10)</td>
<td>7.20e-1</td>
<td>6.80e-1</td>
<td>3.00e-1</td>
<td>3.70e-1</td>
</tr>
</tbody>
</table>

Table 2: Mean and SD of GD indicator obtained by six algorithms.

<table>
<thead>
<tr>
<th>Problem (τₜ, nₜ)</th>
<th>DSSGA-II</th>
<th>PPS</th>
<th>NDSS</th>
<th>DVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDA (25, 10)</td>
<td>3.60e-1</td>
<td>3.00e-1</td>
<td>1.20e-1</td>
<td>1.60e-1</td>
</tr>
<tr>
<td>FDA (30, 10)</td>
<td>4.50e-1</td>
<td>3.80e-1</td>
<td>1.80e-1</td>
<td>2.20e-1</td>
</tr>
<tr>
<td>FDA (35, 10)</td>
<td>5.40e-1</td>
<td>4.80e-1</td>
<td>2.10e-1</td>
<td>2.60e-1</td>
</tr>
<tr>
<td>FDA (40, 10)</td>
<td>6.30e-1</td>
<td>5.80e-1</td>
<td>2.50e-1</td>
<td>3.10e-1</td>
</tr>
<tr>
<td>FDA (45, 10)</td>
<td>7.20e-1</td>
<td>6.80e-1</td>
<td>3.00e-1</td>
<td>3.70e-1</td>
</tr>
</tbody>
</table>

Table 2: Mean and SD of GD indicator obtained by six algorithms.

<table>
<thead>
<tr>
<th>Problem (τₜ, nₜ)</th>
<th>DSSGA-II</th>
<th>PPS</th>
<th>NDSS</th>
<th>DVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDA (25, 10)</td>
<td>3.60e-1</td>
<td>3.00e-1</td>
<td>1.20e-1</td>
<td>1.60e-1</td>
</tr>
<tr>
<td>FDA (30, 10)</td>
<td>4.50e-1</td>
<td>3.80e-1</td>
<td>1.80e-1</td>
<td>2.20e-1</td>
</tr>
<tr>
<td>FDA (35, 10)</td>
<td>5.40e-1</td>
<td>4.80e-1</td>
<td>2.10e-1</td>
<td>2.60e-1</td>
</tr>
<tr>
<td>FDA (40, 10)</td>
<td>6.30e-1</td>
<td>5.80e-1</td>
<td>2.50e-1</td>
<td>3.10e-1</td>
</tr>
<tr>
<td>FDA (45, 10)</td>
<td>7.20e-1</td>
<td>6.80e-1</td>
<td>3.00e-1</td>
<td>3.70e-1</td>
</tr>
</tbody>
</table>

Table 2: Mean and SD of GD indicator obtained by six algorithms.
Table 3: Mean and SD of SP indicator obtained by six algorithms.

<table>
<thead>
<tr>
<th>Prob. (τ1, τ2)</th>
<th>FDA1</th>
<th>DNEGA-II</th>
<th>DDS</th>
<th>BPS</th>
<th>SGIA</th>
<th>DVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>(20, 10)</td>
<td>5.30e-3(9.98e-5)</td>
<td>5.00e-3(1.53e-4)</td>
<td>5.03e-3(3.82e-4)</td>
<td>8.20e-3(1.20e-3)</td>
<td>8.00e-3(1.30e-3)</td>
<td>3.87e-4(9.45e-5)</td>
</tr>
<tr>
<td>(25, 10)</td>
<td>5.46e-3(4.84e-3)</td>
<td>4.85e-3(2.21e-4)</td>
<td>4.10e-3(3.72e-5)</td>
<td>7.80e-3(7.63e-5)</td>
<td>7.60e-3(8.16e-5)</td>
<td>3.57e-4(8.07e-5)</td>
</tr>
<tr>
<td>(30, 10)</td>
<td>5.67e-3(3.18e-3)</td>
<td>4.56e-3(1.46e-4)</td>
<td>3.75e-3(3.64e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
</tr>
<tr>
<td>(30, 10)</td>
<td>5.65e-3(3.18e-3)</td>
<td>4.56e-3(1.46e-4)</td>
<td>3.75e-3(3.64e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
</tr>
<tr>
<td>(30, 10)</td>
<td>5.65e-3(3.18e-3)</td>
<td>4.56e-3(1.46e-4)</td>
<td>3.75e-3(3.64e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
</tr>
</tbody>
</table>

Table 3: Mean and SD of SP indicator obtained by six algorithms.

<table>
<thead>
<tr>
<th>Prob. (τ1, τ2)</th>
<th>FDA1</th>
<th>DNEGA-II</th>
<th>DDS</th>
<th>BPS</th>
<th>SGIA</th>
<th>DVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>(20, 10)</td>
<td>5.30e-3(9.98e-5)</td>
<td>5.00e-3(1.53e-4)</td>
<td>5.03e-3(3.82e-4)</td>
<td>8.20e-3(1.20e-3)</td>
<td>8.00e-3(1.30e-3)</td>
<td>3.87e-4(9.45e-5)</td>
</tr>
<tr>
<td>(25, 10)</td>
<td>5.46e-3(4.84e-3)</td>
<td>4.85e-3(2.21e-4)</td>
<td>4.10e-3(3.72e-5)</td>
<td>7.80e-3(7.63e-5)</td>
<td>7.60e-3(8.16e-5)</td>
<td>3.57e-4(8.07e-5)</td>
</tr>
<tr>
<td>(30, 10)</td>
<td>5.67e-3(3.18e-3)</td>
<td>4.56e-3(1.46e-4)</td>
<td>3.75e-3(3.64e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
</tr>
<tr>
<td>(30, 10)</td>
<td>5.65e-3(3.18e-3)</td>
<td>4.56e-3(1.46e-4)</td>
<td>3.75e-3(3.64e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
</tr>
<tr>
<td>(30, 10)</td>
<td>5.65e-3(3.18e-3)</td>
<td>4.56e-3(1.46e-4)</td>
<td>3.75e-3(3.64e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
</tr>
<tr>
<td>(30, 10)</td>
<td>5.65e-3(3.18e-3)</td>
<td>4.56e-3(1.46e-4)</td>
<td>3.75e-3(3.64e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
<td>3.60e-3(8.60e-5)</td>
</tr>
</tbody>
</table>

5.3. Comparison of the distribution of the final obtained population

In order to analyze the performance of the final population, we randomly selected four test problems from the FDA, dMOP and JY test suites, including FDA1, dMOP2, JY2 and JY6. The final population gained by six algorithms at different stages are shown in figures 7, 8 and 9.

From Figure 7 it is clear that all of the algorithms on FDA1 have better distribution and convergence except for RIS and PPS. This indicates that the RIS strategy alone cannot solve dMOP problems very well. As for PPS, due to the lack of historical information in the early stages, the algorithm cannot converge, but with an increasing amount of information, the distribution and
convergence of the algorithm improves. Unlike FDA1, the POF of dMOP2 changes when the environment changes. We show the POF and population at different times in Fig. 8. We can clearly see that RIS and DNSGA-II cannot solve this kind of problem very well. Similar to FDA1, the performance of PPS at early stages was not good. The distribution of DMS is better than SGEA, but the convergence of SGEA is better than DMS. DVA was superior to SGEA and DMS in both distribution and convergence except when time $t = 10$. The reason is that when $t = 10$, the environment changes dramatically, while DVA does not provide a better initial population, which results in a decrease in the convergence rate of the algorithm. On the JY test suites, as can be seen from Fig. 9 and Fig. 10, DMS and SGEA performed well, but slightly worse than DVA.

6. Discussion

In this section, we mainly discuss the impact of different parts of the algorithm on the whole. The DVA algorithm consists of a response strategy and DEE. The response strategy consists of two main components, one is the prediction part, and the other is the adaptive selection part of the two populations based on the previous prediction part. To deeply examine the role that each component plays in dynamic optimization, we created two variants of the original DVA. The first variant (S1) does not use the adaptive selection which means that the initial population consists entirely of individuals produced by the prediction component. The second variant (S2) contains a complete response strategy, but the static part of the algorithm uses NSGA-II instead of DEE. On the one hand, this is to verify the impact of DEE on the whole; on the other hand, it is also to verify whether our proposed strategy can effectively solve DMOPs when combined with other algorithms. These two variants are compared with the original DVA on four problems with settings of $(τ, n_t) = (20, 10)$. Table 4 presents the average and standard deviations of the metrics obtained by DVA and its variants. On FDA1, JY1 and JY9, it can be clearly seen that s1 is not as good as the original DVA in convergence and distribution, but it
is obviously better than DVA-S2. Combined with Tables 1, 2 and 3 above, DVA containing only the prediction part is still significantly better than other algorithms, which indicates that the DVA prediction part can effectively handle DMOPs. On dMOP1, DVA-S1 has SP and IGD values similar to the original DVA, but the GD of DVA-S1 is significantly smaller than the original DVA. In general, on dDMOP1, DVA-S1 were better than the original DVA and DVA-S2 because the problem of the first two random initializations is too dense on the dMOP1 problem. In order to balance the distribution as much as possible, the original DVA sacrifices the solution with good convergence, and those solutions with relatively good distribution are selected for the initial population. Thus, DVA-S1 is obviously better than the original DVA in convergence. We next analyze the second variant of DVA. Although DVA-S2 is not as good as DVA-S1 and the original DVA in terms of convergence and distribution, DVA-S2 was obviously superior to DNSGA-II under the same static algorithm on dMOP1, JY1 and JY9. This means that our proposed response strategy combined with other related static algorithms can also effectively solve DMOPs.

Table 4: Mean and SD of GD, IGD and SP indicators of DVA variants

<table>
<thead>
<tr>
<th>Prob</th>
<th>Indicator</th>
<th>DVA-S1</th>
<th>DVA-S2</th>
<th>DVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDA1</td>
<td>GD</td>
<td>3.00e-3(9.31e-4)</td>
<td>5.10e-3(2.87e-4)</td>
<td>2.12e-3(1.72e-4)</td>
</tr>
<tr>
<td></td>
<td>IGD</td>
<td>5.18e-3(7.59e-4)</td>
<td>8.22e-3(5.03e-4)</td>
<td>4.48e-3(1.43e-4)</td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>3.74e-3(1.86e-4)</td>
<td>4.93e-3(1.03e-3)</td>
<td>3.66e-3(8.95e-5)</td>
</tr>
<tr>
<td>dMOP1</td>
<td>GD</td>
<td>5.47e-3(7.54e-3)</td>
<td>1.28e-2(5.48e-3)</td>
<td>9.23e-3(1.11e-2)</td>
</tr>
<tr>
<td></td>
<td>IGD</td>
<td>1.48e-2(9.78e-3)</td>
<td>1.89e-2(7.78e-3)</td>
<td>1.52e-2(1.55e-2)</td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>3.45e-3(7.67e-5)</td>
<td>4.04e-3(3.52e-4)</td>
<td>3.55e-3(1.25e-4)</td>
</tr>
<tr>
<td>JY1</td>
<td>GD</td>
<td>2.88e-3(3.32e-4)</td>
<td>4.72e-3(4.55e-4)</td>
<td>2.48e-3(1.04e-3)</td>
</tr>
<tr>
<td></td>
<td>IGD</td>
<td>6.42e-3(2.93e-4)</td>
<td>1.03e-2(6.74e-4)</td>
<td>6.08e-3(1.07e-3)</td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>5.00e-3(1.49e-4)</td>
<td>6.35e-3(4.83e-4)</td>
<td>4.86e-3(3.08e-4)</td>
</tr>
<tr>
<td>JY9</td>
<td>GD</td>
<td>1.28e-2(2.13e-3)</td>
<td>1.51e-2(3.01e-3)</td>
<td>1.00e-2(2.42e-3)</td>
</tr>
<tr>
<td></td>
<td>IGD</td>
<td>1.48e-2(1.82e-3)</td>
<td>2.02e-2(2.10e-3)</td>
<td>1.45e-2(2.18e-3)</td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>6.93e-3(4.76e-4)</td>
<td>9.16e-3(1.86e-3)</td>
<td>6.68e-3(3.50e-4)</td>
</tr>
</tbody>
</table>

In addition, in order to verify the validity of DVA under the same benchmarks, we compared DVA with a dynamic multi-objective evolutionary algorithm based on intensity of environmental change(IEC) [52]. In IEC, the algorithm
effectively tracks the POF according to the intensity of environmental change, and in this paper, JY is also used as the standard test problem set. The specific experimental results are shown in Table 5 and we can intuitively see that in JY1, JY5, JY6, JY7 and JY8, the performance of DVA is significantly better than IEC.

Table 5: Mean and SD of IGD of IEC and DVA

<table>
<thead>
<tr>
<th>Prob (nt, τt)</th>
<th>IEC [52]</th>
<th>DVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>JY1 (10,30)</td>
<td>6.11e-3(4.03e-5)</td>
<td><strong>5.94e-3(4.33e-4)</strong></td>
</tr>
<tr>
<td>JY2 (10,30)</td>
<td>4.98e-2(3.92e-5)</td>
<td><strong>5.04e-2(2.76e-4)</strong></td>
</tr>
<tr>
<td>JY3 (10,30)</td>
<td>3.10e-1(1.64e-3)</td>
<td><strong>3.12e-1(1.76e-3)</strong></td>
</tr>
<tr>
<td>JY4 (10,30)</td>
<td>1.97e-2(6.75e-5)</td>
<td><strong>2.08e-2(3.61e-4)</strong></td>
</tr>
<tr>
<td>JY5 (10,30)</td>
<td>6.52e-3(6.41e-5)</td>
<td><strong>4.22e-3(5.34e-5)</strong></td>
</tr>
<tr>
<td>JY6 (10,30)</td>
<td>4.03e-1(6.38e-2)</td>
<td><strong>3.25e-1(5.31e-2)</strong></td>
</tr>
<tr>
<td>JY7 (10,30)</td>
<td>2.02e+0(4.86e-1)</td>
<td><strong>5.30e-1(4.06e-1)</strong></td>
</tr>
<tr>
<td>JY8 (10,30)</td>
<td>1.51e-2(2.67e-4)</td>
<td><strong>7.52e-3(2.84e-4)</strong></td>
</tr>
<tr>
<td>JY9 (10,30)</td>
<td><strong>6.66e-3(1.19e-4)</strong></td>
<td>1.36e-2(1.28e-3)</td>
</tr>
</tbody>
</table>

7. Conclusions and future work

In order to quickly respond to environmental changes and make the population adapt to the new environment better, we propose DVA to solve DMOPs. DVA as a response strategy guides the whole population to evolve to the next POF when environmental changes are detected. Although the prediction may be inaccurate, dimension-based analysis ensures population diversity. The idea of the DVA algorithm is to analyze the impact of each decision variable on individuals in the current environment, find those dimensions that affect
the distribution, combine with historical information, adopt different prediction methods to generate solutions, and then select those solutions which have good convergence and distribution in the current environment to form the initial population. Compared with five other DMOEAs, experiments show that DVA responds more quickly to environmental changes in those problems with non-linear relationships of decision variables, and it responds better to most of the problems when environmental changes are relatively stable, especially when dealing with complex non-linear problems.

Several extensions are possible for future work:

- Although DVA has great advantages compared with other algorithms, it lacks the correlation analysis between decision variables before and after environmental changes, so it is necessary to design corresponding methods to analyze the correlation between them.

- The DMOEAs have demonstrated the ability to handle constrained D-MOPs and solving constrained DMOPs is a prospective research problem. Simultaneously, new dynamic benchmarks and performance metrics are needed to evaluate the performance of the algorithm.

- DMOEAs should be more applied in practical applications in other fields, such as scheduling [10], control [11] [12] and planning [13] [53].

Acknowledgement

This work was supported by the research projects: the National Natural Science Foundation of China under Grant Nos. 61502408, 61673331, 61772178 and 61403326, the postgraduate research and innovation Project of Hunan Province under Grant No. XDCX2019B057.
References


Figure 5: Evolution curves of average IGD values for seven problems with $n_T = 10$ and $\tau_T = 30$. 
Figure 6: Evolution curves of average IGD values for nine problems with $n_T = 10$ and $\tau_T = 30$. 

(a) JY1  
(b) JY2  
(c) JY3  
(d) JY4  
(e) JY5  
(f) JY6  
(g) JY7  
(h) JY8  
(i) JY9
Figure 7: Solution sets obtained by six algorithms at six different time steps on FDA1.
Figure 8: Solution sets obtained by six algorithms at six different time steps on dMOP2.
Figure 9: Solution sets obtained by seven algorithms at six different time steps on JY2.
Figure 10: Solution sets obtained by seven algorithms at six different time steps on JY6.