Applied Intelligence (2019) 49:1098–1126
https://doi.org/10.1007/s10489-018-1319-7

Multiobjective optimal control for wastewater treatment process using adaptive MOEA/D

Hongbiao Zhou[1,2,3] · **Junfei Qiao**[1,2]

Published online: 25 October 2018

Abstract

Through the analysis of the biological wastewater treatment process (WWTP), a multiobjective optimal control strategy is developed with the usage of energy consumption (EC) and effluent quality (EQ) as objectives to be optimized. To effectively handle the multiobjective optimization problem (MOP) with complex Pareto-optimal front (POF), an adaptive multiobjective evolutionary algorithm based on decomposition (AMOEA/D) is proposed in this paper. Since the efficiency of the multiple reference points and two-phase optimization strategies in solving MOPs with complex POFs has been proved. In the proposed AMOEA/D, an auto-switching strategy based on the aggregation function enhancement is designed to automatically make the algorithm switch from the first phase to the second phase. Besides, an adaptive differential evolution strategy is introduced into AMOEA/D to balance exploration and exploitation during the evolutionary process. Finally, the dynamic optimization, intelligent decision and bottom tracking control of the set-points of the dissolved oxygen and nitrate nitrogen in the WWTP are achieved via the combination of AMOEA/D with the self-organizing fuzzy neural network approximator and the self-organizing fuzzy neural network controller. The international benchmark simulation model No. 1 (BSM1) is utilized for experimental verification. Simulation results demonstrate that the proposed AMOEA/D can effectively reduce the EC of the WWTP under the premise of ensuring effluent parameters to meet the effluent discharge standards.

Keywords Wastewater treatment process · Multiobjective optimal control · MOEA/D · Two-phase optimization · Auto-switching · Adaptive differential evolution strategy

1 Introduction

The activated sludge process is one of the widely used technologies in the wastewater treatment processes (WWTPs) [1, 2]. In this method, activated sludge is formed after a certain reaction time as a result of the propagation of aerobic microorganisms via the continuous filling of air into wastewater. Under the action of bio condensation, adsorption and oxidation of the activated sludge organic pollutants are decomposed and then removed from the wastewater [3]. To meet the effluent discharge standards and reduce fines, wastewater treatment plants are typically operated at the full load status. In this case, the dissolved oxygen concentration (S_O) in the aerobic zone and the nitrate nitrogen level (S_{NO}) in the anaerobic zone are maintained at a high level through aeration and pumping [4]. However the operation of blowers and return sludge pumps requires a large amount of energy supply, resulting in higher operating costs. Besides, according to the biochemical reaction mechanism in the WWTPs, only the appropriate set-points of S_O and S_{NO} can ensure the smooth progress of nitrification and denitrification [5]. Therefore, the set-points of S_O and S_{NO} should be dynamically optimized based on the actual operating conditions so as to reduce energy consumption (EC) and effluent quality (EQ) as much as possible. Note that EQ is utilized to represent fines which are required to be paid due to the discharge of pollutants into the receiving water bodies in the benchmark simulation platform. To improve the treatment effect and reduce the

✉ Junfei Qiao
hyitzhb@163.com

[1] Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

[2] Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China

[3] Faculty of Automation, Huaiyin Institute of Technology, Huai'an 223003, China

running cost, it is imperative to develop an optimal control strategy in the WWTPs.

With the characteristics of nonlinearity, time variation and big lag, the control issues in the WWTPs have been extensively investigated. For example, Wahab et al. designed a multi-variable controller for S_O and S_{NO} by tuning the parameters of PID controller in [6]. Song et al. applied a robust PID controller to the tracking control of S_O in [7] and the results showed that the controller could exhibit good robustness in the case of a model mismatch. Furthermore, Holenda et al. proposed a model predictive controller (MPC) for S_O based on a third-order simplified model of the WWTP in [8] and the results indicated that MPC owns a remarkable performance under the conditions of constant and variable concentrations of S_O. In [9], Belchior et al. investigated the application of the adaptive fuzzy control (AFC) for S_O in which the consequent parameters of fuzzy rules are adaptively adjusted by Lyapunov comprehensive analysis approach. Qiao et al. introduced a self-organizing fuzzy neural network controller for S_O in [10]. For these controllers [6–10], a group of fixed set-points which remains unchanged in the control process is preset. Therefore, the above control strategies could not meet the urgent need of wastewater treatment plants for efficiency promotion, energy saving and consumption reduction.

To effectively reduce EC in the WWTPs, the single-objective optimal control (SOOC) strategy considering EC as the only optimization objective and S_O and S_{NO} as the main decision variables has been proposed [11–14]. For example, Santin et al. presented a two-level hierarchical control structure in [13]. In this method, an upper controller was used to adjust the set-points of S_O online and a lower controller was designed to track the set-points. The results showed that the proposed control strategy can effectively reduce the operating cost. In [14], a genetic algorithm was utilized to optimize EC so that EC can be reduced by dynamically optimizing the set-points of S_O. However, these SOOC methods mainly focus on EC and might easily cause effluent parameters to exceed the standard, resulting in an increase of the total operating cost. Thus, weight factor scheme, which can convert the multiobjective optimization problem (MOP) into a single-objective optimization problem (SOP), is introduced to construct the loss function [15–17]. For example, a data-driven optimal controller (DDAOC) based on the adaptive dynamical programming was introduced to optimize the set-points of S_O and S_{NO} in [16]. The results showed that the proposed DDAOC can provide a reduction of 5.30% in EC. In [17], a Hopfield neural network optimal controller based on the Lagrange multiplier was designed to optimize the set-points of S_O and S_{NO}. In [18], a set-point optimization technique for WWTPs that is composed of dynamic real time optimization and nonlinear model predictive control (RTO-NMPC) was proposed. However, it is hard to obtain a balance between EC and EQ for these methods [16–18], since the determination of a suitable weight factor is difficult.

Fortunately, multiobjective optimization algorithm can overcome the shortcomings described above. Therefore, the design of multiobjective optimal controllers for the WWTPs has been extensively investigated [19–25]. In [20], an interactive software designed based on nondominated sorting genetic algorithm (NSGAII) called IND-NIMBUS, was utilized to construct an optimal operation model for the WWTP. In [21], NSGAII was adopted to establish a multiobjective optimization model for greenhouse emission, operating cost, and effluent contaminant concentration. The results revealed that the improvement of water quality and the decrease of operation cost can cause the increase of greenhouse emission. In [22], a dynamic multiobjective optimization algorithm was applied to establish an optimization model for running cost and wastewater treatment quality. The results demonstrated that these two objectives have conflicting characteristics. Furthermore, some other optimal controllers based on NSGAII for the WWTPs could be found in [23–25]. Recently, an optimal controller based on adaptive fuzzy neural network was developed to solve the problem of high EC in WWTPs, and a good energy-saving effect was achieved [26, 27]. From the analysis of the previous studies [20–27], it is easy to find that the multiobjective optimal controller can not only ensure that the effluent parameters satisfy the standard, but also reduce EC and operation cost effectively. However, there are still many challenges for developing a suitable optimal controller to tackle the multiple conflicting objectives. In addition, NSGAII has a high computational cost, and the convergence, diversity and coverage of the approximated optimal solutions should be enhanced.

The multiobjective evolutionary algorithm based on decomposition (MOEA/D) has been shown to be very efficient in solving MOPs. In MOEA/D, a MOP is decomposed into a number of scalar optimization subproblems and these subproblems are optimized simultaneously in a collaborative manner [28]. Decomposition mechanisms are utilized to push the population to approach the Pareto-optimal front (POF), while a set of uniformly distributed weight vectors are applied to maintain the population diversity [29, 30]. In addition, the concept of subproblem neighborhood, firstly presented in MOEA/D, can improve the balance between exploration and exploitation [31, 32]. However, recent research indicated that MOEA/D can only tackle MOPs with simple POFs, but cannot provide a good distribution for MOPs with irregular POFs. Due to the nonlinearity, big lag and strong interference, the MOPs in the WWTPs may have

complex POFs, such as discontinuity, long tail and/or sharp peak, which could significantly degrade the performance of MOEA/D.

To improve the capability of MOEA/D in addressing MOPs with complex POFs, different approaches have been proposed in recent studies. In [33], Qi et al. proposed an adaptive weight adjustment (AWA) strategy for MOEA/D (MOEA/D-AWA). In MOEA/D-AWA, the AWA strategy is adopted to regularly adjust the distribution of the weight vectors. In [34], Yang et al. investigated the influence of the penalty factor θ in the penalty boundary intersection (PBI) decomposition method on the performance of MOEA/D, thus proposing an MOEA/D with an adaptive penalty scheme (MOEA/D-APS). In [35], Jiang and Yang introduced a kind of MOEA/D with a two-phase strategy and a new niche scheme (MOEA/D-TPN), aiming to find more boundary solutions for complex MOPs. In MOEA/D-TPN, the evolution process is divided into two phases. In the first phase, the ideal point is adopted as the reference point in Tchebycheff decomposition method. In the second phase, the nadir point is utilized as the reference point. When the generations reach a certain number, the algorithm can determine whether or not to execute the second phase based on the crowdedness. In [36], Wang et al. proposed a MOEA/D with multiple reference points (MOEA/D-MR), in which the ideal point and reference point are simultaneously utilized to optimize the subproblems. In [37], Ho-Huu et al. presented an improved MOEA/D (iMOEA/D), in which all the weight vectors are divided into odd and even weight vectors firstly, and then subproblems are optimized by using a two-phase strategy. Simulation results demonstrated the effectiveness of these methods on benchmark testing. However, it is still difficult to apply them to practical engineering. For example, the number of iterations for the first phase still needs to be set by experience in MOEA/D-TPN and tri-objective MOPs cannot be addressed by iMOEA/D

Furthermore, Li and Zhang pointed out that the simulated binary crossover (SBX) operator applied in MOEA/D often generates inferior solutions in [38]. Therefore, they proposed a well-known MOEA/D with a differential evolution (DE) operator, called MOEAD-DE. Considering the defects of a single DE operator, Li et al. applied four DE operators to form an operator pool and designed a bandit-based adaptive operator selection strategy which significantly improved the performance of the MOEA/D [39]. The DE operator pool strategy was also utilized in [40, 41] and [42]. These promising results encourage us to design a suitable operator selection strategy to enhance the proposed improved MOEA/D.

In this paper, to effectively address the MOPs in the WWTPs an adaptive MOEA/D algorithm (AMOEA/D) is proposed. In AMOEA/D, the advantages of multiple reference points and two-phase optimization in MOEA/D-TPN are still retained. To overcome the shortcomings of iteration setting for the first phase by experience, an auto-switching scheme which could automatically switch the algorithm to the second phase is designed based on the aggregation function enhancement (AFE). In addition, an adaptive operator selection (AOS) scheme, which can select appropriate DE operator online, is developed based on the solution replacement rate. Furthermore, a hybrid multiobjective optimal control (HMOOC) strategy is designed to reduce EC without violating effluent standards. First, a self-organizing fuzzy neural network approximator developed in the previous study [43, 44] is utilized to establish the objective functions of EC and EQ, according to the analysis of the control variables and optimization objectives. Second, the proposed AMOEA/D in this study is adopted to dynamically optimize the set-points of S_O and S_{NO}. Then, an intelligent decision-making system based on the fuzzy membership function method is applied to select the preferred solution as the optimized set-points of the current optimization cycle. Third, a self-organizing fuzzy neural network controller designed in the previous study [45] is used to track the optimal set-points. Finally, the international benchmark simulation model No.1 (BSM1) is introduced to validate the effectiveness of the proposed AMOEA/D-based HMOOC strategy.

The rest of this paper is organized as follows. Section 2 introduces the MOP in the WWTP. In Section 3, the proposed AMOEA/D is described in detail. Section 4 presents the HMOOC strategy for WWTP. Experimental studies for AMOEA/D on benchmark test instances and WWTP are presented in Sections 5 and 6, respectively. Finally, Section 7 concludes this paper.

2 MOP in the WWTP

2.1 BSM1

WWTP is a large nonlinear dynamic systems subject to large perturbations in influent flow rate and pollutant load, together with uncertainties concerning the composition of the incoming wastewater. To evaluate the possible control strategies, a benchmark was proposed by the Working Groups of COST Action 682 and 624-BSM1 [46]. The general overview of the BSM1 plant is depicted in Fig. 1, where the plant is composed of the biological reactor and the secondary settler. The biological reactor consists of five tanks connected in cascade. Tanks 1 and 2 are non-aerated but fully mixed with volume equal to 1000 m^3 each. Tanks 3, 4 and 5 are aerated and their volumes are approximately equal to 1333 m^3 each. The activated sludge model No.1 (ASM1) is selected to describe the biological

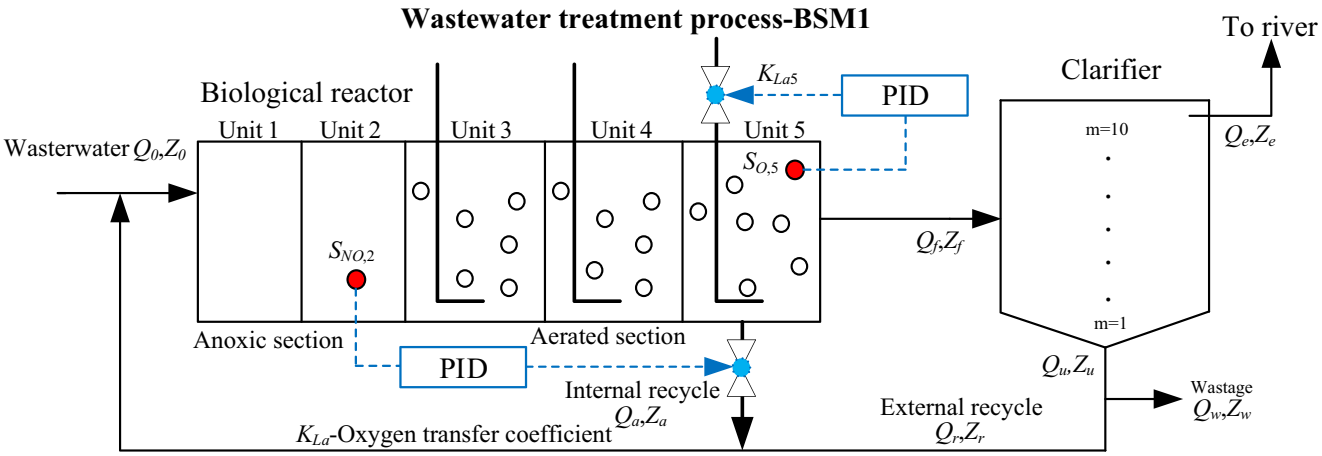

Fig. 1 General overview of the BSM1 plant

phenomena taking place in the biological reactor. In BSM1, there are two control loops, namely the S_O and S_{NO} control loops. The first control loop tunes the dissolved oxygen concentration in the fifth tank $S_{O,5}$ by manipulating the oxygen transfer coefficient K_{La5}, and the second one tunes the nitrate nitrogen level in the second tank $S_{NO,2}$ by manipulating the internal recirculation flow rate Q_a.

The general equations for mass balancing in BSM1 are as follows [46]:

For unit 1(k =1):

$$\frac{dZ_1}{dt} = \frac{1}{V_1}(Q_a Z_a + Q_r Z_r + Q_0 Z_0 + r_1 V_1 - Q_1 Z_1) \quad (1)$$

$$Q_1 = Q_a + Q_r + Q_0 \quad (2)$$

where Q_a, Q_r, Q, and Q_1 are, respectively, the internal recirculation flow rate, the external recirculation flow rate, the influent flow rate, and the flow rate in unit 1; Z_a, Z_r, Z_0, and Z_1 are, respectively, the component concentration of the internal recirculation, the external recirculation, the influent, and unit 1; r_1 is the component reaction rate in unit 1 and V_1 is the volume of unit 1.

For unit 2-5 (k =2-5):

$$\frac{dZ_k}{dt} = \frac{1}{V_k}(Q_{k-1} Z_{k-1} + r_k V_k - Q_k Z_k) \quad (3)$$

$$Q_k = Q_{k-1} \quad (4)$$

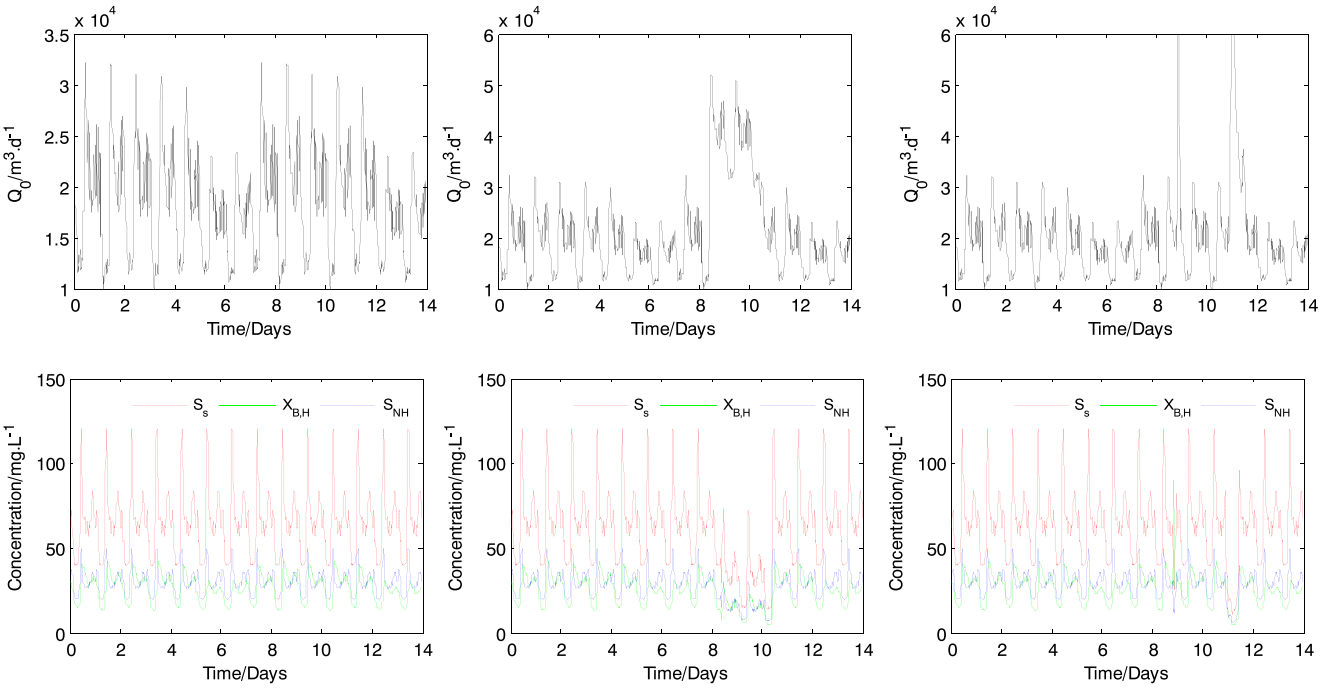

Fig. 2 The influent and S_S, $X_{B,H}$, S_{NH} concentrations in three working conditions

where Q_k, Z_k, and r_k are, respectively, the flow rate, the component concentration, and the component reaction rate in unit k. V_k is the volume of the kth unit.

In BSM1, three influent files, including dry weather, rain weather, and storm weather, are generated from a practical wastewater treatment plant. The data sampling interval is 15 min [46]. The plots of the influent flow rates and the influent S_S, X_{BH} and S_{NH} concentrations under three different weather situations are presented in Fig. 2, which can indicate the characteristics of strong nonlinearity, severe uncertainty and strong coupling in the WWTP [46, 47].

2.2 MOP in the WWTP

The aim of multiobjective optimal control is to achieve the best balance between EC and EQ by dynamically adjusting the set-points of $S_{O,5}$ and $S_{NO,2}$. In the WWTP, the set-points of $S_{O,5}$ and $S_{NO,2}$ not only affect EC but also show a close relationship with EQ. In this study, the proposed AMOEA/D is utilized to address these two conflicting objectives simultaneously.

Since the aeration energy consumption (AE) and the pump energy consumption (PE) account for more than 70% of the total energy consumption, the optimization objective EC is defined as the sum of AE and PE [24], as follows:

$$EC = AE + PE \tag{5}$$

As defined in BSM1, AE and PE can be calculated as [46]

$$AE = \frac{S_{O,sat}}{T \cdot 1.8 \cdot 1000} \int_{kT}^{(k+1)T} \sum_{i=1}^{5} V_i \cdot K_{Lai}(t)dt \tag{6}$$

$$PE = \frac{1}{T \cdot 1000} \int_{kT}^{(k+1)T} (4Q_a(t) + 8Q_r(t) + 50Q_w(t))dt \tag{7}$$

where V_i and K_{Lai} are the volume and oxygen transfer coefficient of the ith unit, respectively. $S_{O,sat}$ is the saturation concentration for oxygen. T is the optimal cycle. Q_a, Q_r, and Q_w are, respectively, the internal recirculation flow rate, the external recirculation flow rate, and the sludge flow rate.

Moreover, EQ is defined as

$$EQ = \frac{1}{T \cdot 1000} \int_{t_0}^{t_f} \begin{pmatrix} 2 \cdot SS_e(t) + COD_e(t) \\ +3 \cdot S_{NKj,e}(t) + 10 \cdot S_{NO,e}(t) \\ +2 \cdot BOD_{5,e}(t) \end{pmatrix} Q_e(t)dt \tag{8}$$

where SS_e, COD_e, $S_{NKj,e}$, $S_{NO,e}$, and $BOD_{5,e}$ are, respectively, the effluent concentrations of suspended solid, chemical oxygen demand, Kjeldahl nitrogen, nitrate nitrogen, and biochemical oxygen demand of 5 days; Q_e is the effluent flow rate.

The constraint condition of the MOP is the standard values of five kinds of effluent parameters given in BSM1, as follows [46]:

$$\begin{aligned} &S_{NH,e} \le 4mg/l, N_{tot,e} \le 18mg/l, \\ &BOD_{5,e} \le 10mg/l, COD_e \le 100mg/l, \\ &SS_e \le 30mg/l \end{aligned} \tag{9}$$

where $N_{tot,e}$ is the effluent total nitrogen which is the sum of $S_{NO,e}$ and $S_{NKj,e}$.

To sum up, the MOP is described as follows:

$$\begin{aligned} \text{minimize} \quad & \mathbf{F}(\mathbf{x}) = (f_{EC}(\mathbf{x}), f_{EQ}(\mathbf{x}))^T \\ \text{subject to} \quad & g_j(\mathbf{x}) = f_j(\mathbf{x}) - c_j \le 0, j = 1, 2, ..., 5 \\ & x_i^l \le x_i \le x_i^u, i = 1, 2, ..., 5 \end{aligned} \tag{10}$$

where $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)^T = [Q_w, K_{La3}, K_{La4}, S_{O,5}, S_{NO,2}]^T$ is the decision vector, $g_j(\mathbf{x}) \le 0$ are inequality constraints, $f_j(\mathbf{x})$ represents the relationship between the jth effluent parameter and the decision vector, $xu\ i$ and $xl\ i$ are the upper and lower limits of the ith decision variable.

3 The AMOEA/D algorithm

This section first provides some background knowledge, including the basic definition of constrained multiobjective optimization problem (CMOP) and the Tchebycheff decomposition method. Then, the advantages of multiple reference points and multiple DE operators are described. Furthermore, the AOS-based adaptive DE strategy and the auto-switching-based adaptive two-phase strategy are introduced. Finally, the procedure of the proposed AMOEA/D is given.

3.1 Background

3.1.1 Constrained multiobjective optimization problem

Without loss of generality, this paper considers the following CMOP [48]:

$$\begin{aligned} \text{minimize} \quad & \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), ..., f_m(\mathbf{x}))^T \\ \text{subject to} \quad & g_j(\mathbf{x}) \le 0, j = 1, ..., J \\ & h_k(\mathbf{x}) = 0, k = 1, ..., K \end{aligned} \tag{11}$$

where $\mathbf{x} = (x_1,...,x_n)^T \in \Omega$ is a decision variable vector, Ω is the feasible search region, $\mathbf{F}$: $\Omega \rightarrow R^m$ consists of m real-valued objective functions, R^m is the objective space, $g_j(\mathbf{x}) \le 0$ are inequality constraints, and $h_k(\mathbf{x}) = 0$ are equality constraints.

3.1.2 The Tchebycheff decomposition method

The decomposition method used in the MOEA/D is the Tchebycheff method. In this method, the scalar objective optimization problem is in the form

$$\begin{aligned} &\text{minimize } g^{te}(\mathbf{x}\,|\mathbf{w}\,,\mathbf{z}^*) = \max_{1\le j\le m}\{w_i|f_j(\mathbf{x})-z_j^*|\} \\ &\text{subject to } \mathbf{x}\in\Omega \end{aligned} \tag{12}$$

where $\mathbf{w}=(w_1,...,w_m)^T$ is a weight vector, i.e., $\lambda_j \ge 0$ for all $j=1,...,m$ and $\sum_{j=1}^{m} w_j = 1$. $\mathbf{z}^* = (z_1^*,...,z_m^*)$ is the ideal point in the objective space, i.e., $z_j^* = \min\{f_j(\mathbf{x})|\ \mathbf{x}\in\Omega\}$ for each $j=1,...,m$. Since $\mathbf{z}^*$ is generally unknown before searching, z_j^* can be replaced by $f_j(\mathbf{x})$ with the smallest value during the searching process [28].

3.2 Multiple reference points and multiple DE operators

3.2.1 Multiple reference points

Figure 3 shows the distribution of optimal solutions over a POF by using the ideal point $\mathbf{z}^*$ and the nadir point $\mathbf{z}^{nad}$. It is clear from Fig. 3 that, for a convex MOP, the optimal solutions would converge to the middle region of the POF, if $\mathbf{z}^*$ is used as the reference point. On the contrary, in the case of using $\mathbf{z}^{nad}$ as the reference point, more boundary optimal solutions would be obtained. In addition, for a concave MOP, uniformly distributed optimal solutions can be obtained by using $\mathbf{z}^*$ as the reference point. Otherwise, the optimal solutions would converge to the middle region of the POF, if $\mathbf{z}^{nad}$ is utilized as the reference point. Considering the complementarity of $\mathbf{z}^*$ and $\mathbf{z}^{nad}$, a two-phase strategy was proposed in [35], and a multiple reference points strategy was introduced in [36].

As demonstrated in [35] and [36], the scalar optimization problem using $\mathbf{z}^{nad}$ as reference point can be formed as

$$\begin{aligned} &\text{maxmize } g^{te}(\mathbf{x}\,|\mathbf{w}\,,\mathbf{z}^{nad}) = \min_{1\le j\le m}\{w_j|z_j^{nad}-f_j(\mathbf{x})|\} \\ &\text{subject to } \mathbf{x}\in\Omega \end{aligned} \tag{13}$$

where $\mathbf{z}^{nad} = (z_1^{nad},...,z_m^{nad})$ is the nadir point generated from the worst objective values of the obtained POF, i.e., $z_j^{nad} = \max\{f_j(\mathbf{x})|\mathbf{x}\in\Omega_{\mathbf{x}}\}$ for each $j=1,...,m$.

3.2.2 Multiple DE operators

Due to different DE operators showing different search characteristics, the performance of MOEA/D could be improved through effective combination of them during the evolution process [39]. In this paper, three well-known DE operators (i.e., rand/1/bin, rand/2/bin and rand/1/bin˜, respectively named DE_1, DE_2, and DE_3) with fixed parameters settings are utilized [41]. Their expressions are as follows:

$$DE_1 : u_j^i = \begin{cases} x_j^{r_1} + F\times(x_j^{r_2}-x_j^{r_3}) + F\times(x_j^{r_4}-x_j^{r_5}) & \text{if rand} < CR \\ x_j^i & \text{otherwise} \end{cases}, \text{with } CR = 0.9 \text{ and } F = 0.7 \tag{14}$$

$$DE_2 : u_j^i = \begin{cases} x_j^{r_1} + F\times(x_j^{r_2}-x_j^{r_3}) & \text{if rand} < CR \\ x_j^i & \text{otherwise} \end{cases}, \text{with } CR = 0.5 \text{ and } F = 0.5 \tag{15}$$

$$DE_3 : u_j^i = \begin{cases} x_j^i + F\times(x_j^{r_1}-x_j^{r_2}) & \text{if rand} < CR \\ x_j^i & \text{otherwise} \end{cases}, \text{with } CR = 0.1 \text{ and } F = 0.5 \tag{16}$$

where u_j^i is the jth dimension of the ith trial solution, x_j^i is the jth dimension of the ith parent solution, rand is a random number generated from [0,1], CR is the crossover rate, F is the scaling factor, and r_1, r_2, r_3, r_4, and r_5 are five distinct individuals randomly selected from parent population.

For these three DE operators, the parent solutions of DE_1 and DE_2 are randomly selected from the population. Thus, they have a strong exploration ability and are suitable for the early evolution stage. Compared with DE_2, DE_1 with two random vectors, which could generate larger disturbances, has stronger capability to jump out of local POF, and is suitable for MOPs with complicated POFs. Furthermore, DE_3 could improve local search ability since it can inherit much information from its parent.

In addition, the setting of control parameters CR and F also significantly influences the performance of DE. A large value of CR shows that the offspring solution could inherit

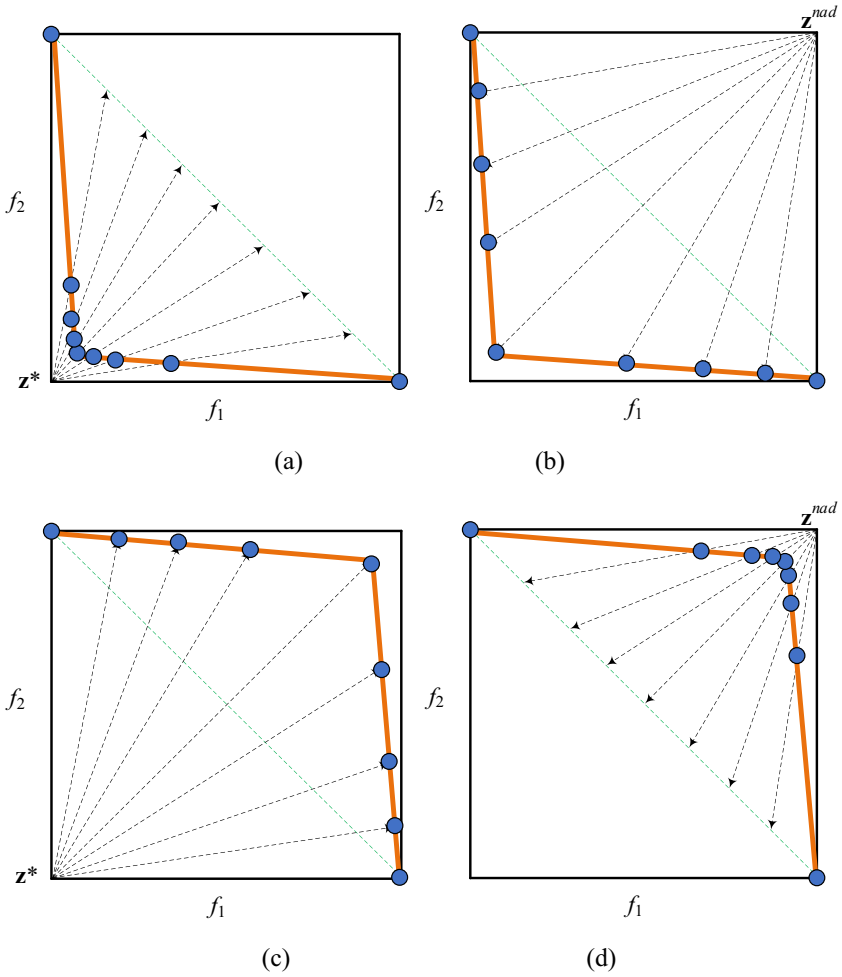

Fig. 3 Pareto optimal solutions obtained by using $\mathbf{z}^*$ and $\mathbf{z}^{nad}$

a mutant gene from the parent with a large probability, improving the search ability around the target vector. Conversely, a small value of CR indicates that the offspring solution could inherit original gene from the parent with a high probability, increasing the exploitation ability around the target vector. For parameter F, a large value of F could increase the search step size, expand the search scope, and improve the population diversity. By contrast, a small value of F could enhance the exploitation ability around the target vector, and then improve the population convergence [41].

3.3 The AMOEA/D algorithm

The proposed AMOEA/D is developed based on MOEA/D-DE [38], which is a well-known improved version of MOEA/D. The flowchart of two-phase optimization with multiple adaptive strategies is first presented. Then, the initialization, AOS-based reproduction operation, constraint handling technique, replacement operation, and auto-switching scheme are briefly introduced. Finally, the pseudo-code of the proposed AMOEA/D is given.

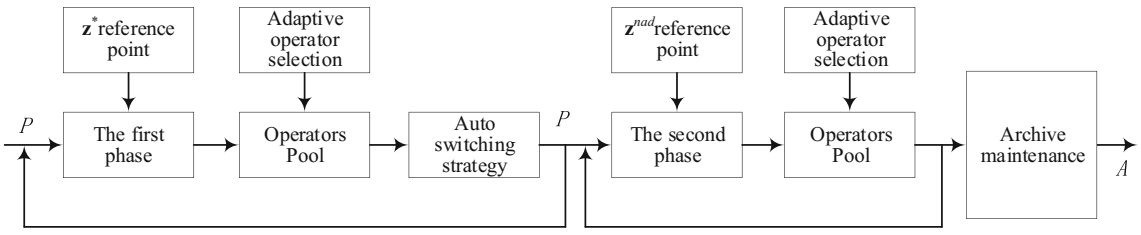

Fig. 4 The flowchart of two-phase optimization with multiple adaptive strategies

Figure 4 shows the flowchart of two-phase optimization with multiple adaptive strategies. Since the concavity-convexity of an MOP is unknown in advance, $\mathbf{z}^*$ is first utilized as the reference point in the first phase. At each iteration, a suitable DE operator selected by the AOS scheme is adopted to produce the offspring solutions. The algorithm runs until the auto-switching scheme detects that it is in a stagnation state, and then automatically switches to the second phase. At the end of the first phase, the obtained optimal solutions are saved to the external archive EA_1. In the second phase, the reference point $\mathbf{z}^{nad}$ is derived from the maximum values of the objective functions at the end of the first phase. Once the preset maximum number of iteration is reached, the algorithm stops and the obtained optimal solutions in the second phase are saved to the external archive EA_2. Finally, the nondominated solutions are identified from EA_1 and EA_2 by external archive maintenance algorithm.

3.3.1 Initialization

In MOEA/D, the initialization operation mainly includes the initialization of population and weight vectors. Since we have no prior knowledge about the position of the POS, the initial population is randomly generated from the decision space. During evolution, the optimization of each subproblem is completed by evolutionary operation between the subproblem and subproblems in its neighborhood. The relationship of adjacent subproblems is determined by the distance between weight vectors associated with subproblems. To a certain extent, the uniform distribution of weight vectors can improve the uniformity of the approximated POF obtained by the algorithm. MOEA/D adopts the simplex lattice design method, proposed by Scheff in 1958, to set weight vectors, as follows:

$$w_j^i \in \left\{ \frac{0}{H}, \frac{1}{H}, \frac{2}{H}, ..., \frac{H}{H} \right\} \tag{17}$$

$$\sum_{j=1}^{m} w_j^i = 1 \tag{18}$$

where each subproblem $i (i = 1,2,...,N)$ corresponds to a weight vector $\mathbf{w}^i = (w_1^i w_2^i,...,w_m^i)$, $\mathbf{w}_j^i \geq 0$, $j = 1,2,...,m$, the weight vector set is $W = (\mathbf{w}^1, \mathbf{w}^2,...,\mathbf{w}^N)$, where $N = C_{H+m-1}^{m-1}$ is the total number of weight vectors [29]. In MOEA/D, a neighborhood of weight vector $\mathbf{w}^i$ is defined as a set of its several closest weight vectors in $\{\mathbf{w}^1, \mathbf{w}^2,...,\mathbf{w}^N\}$. The neighborhood of the ith subproblem consists of all the subproblems with the weight vectors from the neighborhood of $\mathbf{w}^i$. In initialization, we need to compute the Euclidean distances between any two weight vectors and then work out the T closet weight vectors to each weight vector [28]. For each $i = 1, 2,...,N$, set $B(i) = \{i_1, i_2,...,i_T\}$, where $i_1, i_2,...,i_T$ are the T closest weight vectors to $\mathbf{w}^i$. The pseudo-code of initialization is given in Algorithm 1. It is worth noting that the weight vectors need to be reinitialized at the beginning of the second phase, as described in [35].

Algorithm 1 Initialization

```
Generate a set of weight vector W ←{w^1,...,w^N} according to the simplex lattice design method;
Compute the Euclidean distances between any two weight vectors;
for i ←1 to N do
    B(i) ←{i_1,...,i_T} where w^{i1} w^{i2},...,w^{iT} are the T closest weight vectors to w^i;
end
Sample uniformly randomly an initial population P ←{x^1,...,x^N} from the search region;
Evaluate each x^i in P;
Initialize the ideal and reference objective vectors z* and z^r;
```

3.3.2 AOS-based reproduction operation

In MOEA/D-DE, to maintain the population diversity, the maximum number of parent solutions replaced by an offspring solution is bounded by n_r. In [38], the recommended setting of n_r is 2. Let n_p be the maximum number of parent solutions replaced by an offspring solutions for all subproblems at each iteration, the parent solution updating rate at the tth iteration is defined as

$$SRR(t) = \frac{n_p}{n_{p,\max} - n_{p,\min}} \tag{19}$$

where $n_{p,\max}$ is the maximum of n_p ($n_{p,\max} = 2 \times N$), and $n_{p,\min}$ is the minimum of n_p ($n_{p,\min} = 0 \times N$). A large value of n_p indicates that the algorithm is in diversity status, and it is advisable to select DE_1 with large probability p_1 to strengthen the global exploration ability. A small value of n_p shows that the algorithm is in convergence status, and DE_3 operator should be applied with a large probability p_3 to enhance the local exploitation capability. For median n_p, DE_2 operator should be utilized to balance exploration and exploitation. Therefore, as the number of n_p decreases, p_3 should decrease while p_1 increase. In this paper, the expressions of p_1 and p_3 can be defined as follows:

$$p_1(t) = 0.9 \times \frac{1.0}{1 + \exp(-20 \times (SRR(t) - 0.35))} \tag{20}$$

$$p_3(t) = 0.9 \times \frac{1.0}{1 + \exp(20 \times (SRR(t) - 0.2))} \tag{21}$$

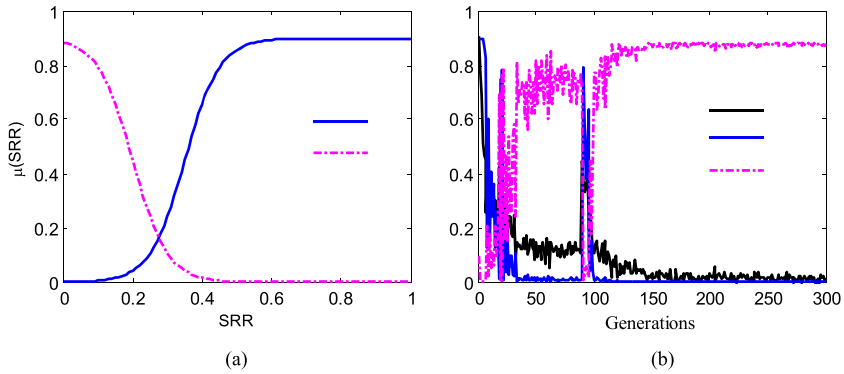

Fig. 5 The relationship and dynamic trend of p_1, p_3 and *SRR* on F1

Figure 5a shows the relationship between p_1, p_3 and *SRR*. Figure 5b shows the dynamic trend of p_1, p_3 and *SRR* on F1 test instance. As seen from Fig. 5, in the initial evolution of the first phase n_p is large, and the DE_1 operator is selected with a large probability to enhance the population diversity. With the operation of the first phase, n_p decreases, and the DE_3 operator is selected with a large probability to improve the local search capability. The same phenomenon could be observed in the second phase. And the pseudo-code of the AOS scheme is presented in Algorithm 2.

Algorithm 2 Adaptive operator selection

```
Record the value of n_p in the current iteration step;
Calculate the value p_1 and p_3 by (20) and (21);
Randomly generate the value of rand in [0,1];
if rand < p_1
    Select DE_1 operator;
else if rand < p_3
    Select DE_3 operator;
else
    Select DE_2 operator;
end
```

After the operation of the DE operator, AMOEA/D also needs to perform the mutation operator. For simplicity, the polynomial mutation operator [29, 38] is applied in AMOEA/D, as follows:

$$x_j'^i = \begin{cases} u_j^i + \sigma_j \times (b_j - a_j) & \text{if rand} < p_m \\ u_j^i & \text{otherwise} \end{cases} \tag{22}$$

$$\sigma_j = \begin{cases} (2 \times rand)^{\frac{1}{\eta+1}} - 1 & \text{if rand} < 0.5 \\ 1 - (2 - 2 \times rand)^{\frac{1}{\eta+1}} & \text{otherwise} \end{cases} \tag{23}$$

where $j \in \{1,2,...,n\}$, rand$\in[0,1]$, p_m is the mutation probability, η is the distribution index, a_j and b_j are the lower and upper bounds of the jth decision variable, respectively. To maintain the population diversity in AMOEA/D, five parent solutions are selected from the whole population with a low probability 1-δ. In such a way, a very wide range of child solutions could be generated due to the dissimilarity among these parent solutions. Therefore, the exploration ability of the search could be enhanced. The pseudo-code of reproduction operation with the AOS scheme is given in Algorithm 3.

Algorithm 3 Reproduction

```
Generate uniformly randomly a number rand from (0,1);
if rand < δ then
    E ← B(i);
else
    E ← P;
end
Select the current solution x^i and randomly select five solutions x^r1, x^r2, x^r3, x^r4, and x^r5 from E;
Select a DE operator according to Algorithm 2;
Generate an offspring u by using the selected DE operator;
Produce a new solution x' by using the PM operator;
Repair x' if an element of x' is out of the boundary;
```

3.3.3 Replacement operation with constraint handling technique

When performing replacement operation, the processing of constraint conditions should be taken into consideration. In this paper, the constraint handing technique introduced

in [48] is utilized. For infeasible solution $\mathbf{x}$ its constraint violation can be defined as

$$cv(\mathbf{x}) = \sum_{j=1}^{J} \max(g_j(\mathbf{x}), 0) + \sum_{j=J+1}^{K} \max(|h_j(\mathbf{x})| - \delta, 0) \quad (24)$$

For parent solution $\mathbf{x}$ and offspring solution $\mathbf{x}'$, $\mathbf{x}$ is substituted by $\mathbf{x}'$ if one of the following conditions is met: (1) Both $\mathbf{x}'$ and $\mathbf{x}$ are feasible solutions, and the aggregation function value of $\mathbf{x}'$ is smaller than that of $\mathbf{x}$; (2) $\mathbf{x}'$ is a feasible solution, and $\mathbf{x}$ is in the infeasible region; (3) Both $\mathbf{x}'$ and $\mathbf{x}$ are in the infeasible region, and the $cv(\mathbf{x}')$ is smaller than $cv(\mathbf{x})$. The pseudo-code of the replacement operation with constraint handling technique is presented in Algorithm 4. It should be noted that the condition of $g(\mathbf{x}'|\mathbf{w}^j\mathbf{z}^*) \leq g(\mathbf{x}^j|\mathbf{w}^j\mathbf{z}^*)$ in the first phase should be changed as $g(\mathbf{x}'|\mathbf{w}^j\ \mathbf{z}^{\text{nad}}) \geq g(\mathbf{x}^j|\mathbf{w}^j\ \mathbf{z}^{\text{nad}})$ in the second phase.

Algorithm 4 Replacement

```
1  Set c =0;
2  while c<n_r && E is not null do
3     Randomly pick an index j from E;
4     Calculate the g value using TCH method for w^j;
5     if cv(x') = cv(x^j) = 0 then
6        if g(x'|w^j z*) ≤ g(x^j|w^j z*) then
7           Set x^j = x', FV^j =FV(x'), c = c+ 1, and
            n_p = n_p+1;
8        end
9     else if cv(x') = 0 & cv(x^j) >0 then
10       Set x^j = x', FV^j =FV(x'), c = c+ 1, and
         n_p = n_p+1;
11    else if cv(x') >0 & cv(x^j) >0 then
12       if cv(x') <cv(x^j) then
13          Set x^j = x', FV^j =FV(x'), c = c+ 1, and
            n_p = n_p+1;
14       end
15    end
16    Delete j from E;
17 end
```

3.3.4 Auto-switching scheme

In [35], a fixed number of iterations is needed to be set in advance for phase switching resulting in an unequal distribution of computing resources. In this paper, an auto-switching scheme is designed based on the aggregation function enhancement (AFE). According to (12), MOEA/D approaches POF by minimizing the aggregation function values of all the subproblems. If the value of the aggregation function of all subproblems could not be improved, it shows that the algorithm has found real POF or is in the stagnation status. At this point, the algorithm should be switched to the second phase from the first phase. The value of *AFE* for the ith subproblem at the tth iteration is calculated as below:

$$AFE^i(t) = g^{te}(w^i, t) - g^{te}(w^i, t-1) \quad (25)$$

where $g^{te}(w^i, t)$ and $g^{te}(w^i, t-1)$ are the aggregation function values of the ith subproblem at the tth and (t-1)th iterations, respectively. Let the maximum of *AFE* for all subproblems at the tth iteration be

$$AFE_{\max}(t) = \max_{1 \leq i \leq N} AFE^i(t) \quad (26)$$

Let *varAFE* be the variance of the vector of $[AFE_{\max}(t-\gamma+1), AFE_{\max}(t-\gamma+2), ..., AFE_{\max}(t)]$, if $varAFE \leq \varepsilon$, the algorithm switches from the first phase to the second phase, where ε is a predefined stopping threshold. In the experiments, the value of ε is selected as $\varepsilon = 10^{-6}$ by trial-and-error.

3.3.5 The procedure of the AMOEA/D algorithm

The pseudo-code of the proposed AMOEA/D is provided in Algorithm 5. At the beginning of AMOEA/D some related parameters (i.e. N, *maxIteration*, p_1, p_3, r, and *flag*) are initialized in line 1. Then, in line 2, the weight vectors W and population P are generated according to Algorithm 1. From line 3 to line 34, the algorithm enters the major circle until the maximum number of iteration is reached and then the algorithm stops. From line 5 to line 20, the algorithm performs the first phase. From line 13 to line 19, the auto-switching scheme is run to determine whether to switch to the second phase. From line 21 to line 33, the algorithm executes the second phase. Besides, it is vital to note that the algorithm needs to store the nondominated solutions obtained in the first phase to the external archive EA_1, reset the reference point with $\mathbf{z}^{\text{nad}}$, reinitialize the weight vectors and recalculate the neighborhood relationship from line 21 to line 24. In line 35, the algorithm needs to save the non-dominated solutions acquired in the second phase to the external archive EA_2 at the end of the second phase. Finally, in line 36, a certain number of non-dominated solutions with good diversity are identified from EA_1 and EA_2 by using the external archive maintenance algorithm with dynamic crowding distance method [49].

Algorithm 5 AMOEA/D

1 Initialize N, *maxIteration*, p_1 =0.9, p_3 =0.1, *flag*=1, γ =10, $\varepsilon = 10^{-6}$;
2 Initialize the population P and weight vectors W according to **Algorithm 1**;
3 **for** t =1 to *maxIteration* **do**
4 Set n_p =0;
5 **if** flag==1
6 **for** i ←1 to N **do**
7 Generate an offspring $\mathbf{x}'$ according to **Algorithm 3**;
8 Evaluate the **F**-function value FV of $\mathbf{x}'$;
9 Compute the degree of constraint violation $cv(\mathbf{x}^i)$ for individual $\mathbf{x}^i$ by (23);
10 Update the current ideal objective vector $\mathbf{z}^*$;
11 Update neighboring solutions according to **Algorithm 4**;
12 **end**
13 Calculate the value of $AFE_{\max}$ by (25) and (26);
14 **if** rem(t, γ) ==0
15 Calculate the variance *varAFE*;
16 **if** *varAFE*≤ ε
17 $flag = 2, flag\_t = t$;
18 **end**
19 **end**
20 **else**
21 **if** $t = flag\_t$
22 Save the population in EA_1;
23 Set the nadir point $\mathbf{z}^{nad} = (z_1^{nad},...,z_m^{nad})$, where z_j^{nad} =max$\{f_j(\mathbf{x})\}$;
24 Re_initialize weight vectors and re_calculate neighborhood relationship;
25 **else**
26 **for** i ←1 to N **do**
27 Generate an offspring $\mathbf{x}'$ according to **Algorithm 3**;
28 Evaluate the **F**-function value FV of $\mathbf{x}'$;
29 Compute the degree of constraint violation $cv(\mathbf{x}^i)$ for individual $\mathbf{x}^i$ by (23);
30 Update neighboring solutions according to **Algorithm 4** with reference point $\mathbf{z}^{nad}$;
31 **end**
32 **end**
33 **end**
34 **end**
35 Save the final population in EA_2;
36 Output the non-dominated solutions from $EA_1 \cup EA_2$;

3.4 Computational cost of one generation of AMOEA/D

In this section, let us consider the computational cost of AMOEA/D in one generation. AMOEA/D has the same framework as MOEA/D-DE [38], thus the increased computation cost is attributed to its detection step for the auto-switching strategy and calculation of the parent solution updating rate for adaptive DE strategy. The calculation of the $AFE_{\max}$ value (line 13 in Algorithm 5) requires $O(mN)$ computations. The calculation of the variance *varAFE* has a smaller computational cost. Therefore, the overall complexity of detection step is $O(mN)$. Consider the adaptive operator selection presented in Algorithm 2, the calculation of the parent solution updating rate (line 2 in Algorithm 2) costs $O(N)$ computations. At last, the reinitialization of the weight vectors and the recalculation of the neighborhood (line 24 in Algorithm 5) require $O(mN)$ and $O(mNT)$ computations, respectively. However, these two operations only need to be run once during the entire evolutionary process. In summary, the computational cost of AMOEA/D in each generation is $O(mN)$.

4 Multiobjective optimal control based on AMOEA/D

By analyzing the (5)–(8), it can be found that EC and EQ have no explicit mathematical relationship with the five decision variables, particularly with $S_{O,5}$ and $S_{NO,2}$. Furthermore, the critical effluent quality parameters forming the constraint condition cannot be measured online. Therefore, the HMOOC strategy is proposed for the WWTP. First, a data-driven-based modeling approach is utilized to establish the accurate soft-computing models for EC, EQ and effluent quality parameters as optimization objectives and constraints. Second, the proposed AMOEA/D, utilizing multiple DE operators with AOS scheme and two-phase optimization with auto-switching scheme, is applied to dynamically optimize $S_{O,5}$ and $S_{NO,2}$ by minimizing the established objectives. Furthermore, an intelligent decision system is developed to select a preferred solution from the obtained POS as an optimized set-points. Finally, the multivariable controller, employing the self-organizing fuzzy neural network, is adopted to track the optimized set-points. And the overall structure of the HMOOC for the WWTP is shown in Fig. 6.

The overall process of the HMOOC for the WWTP is described as follows:

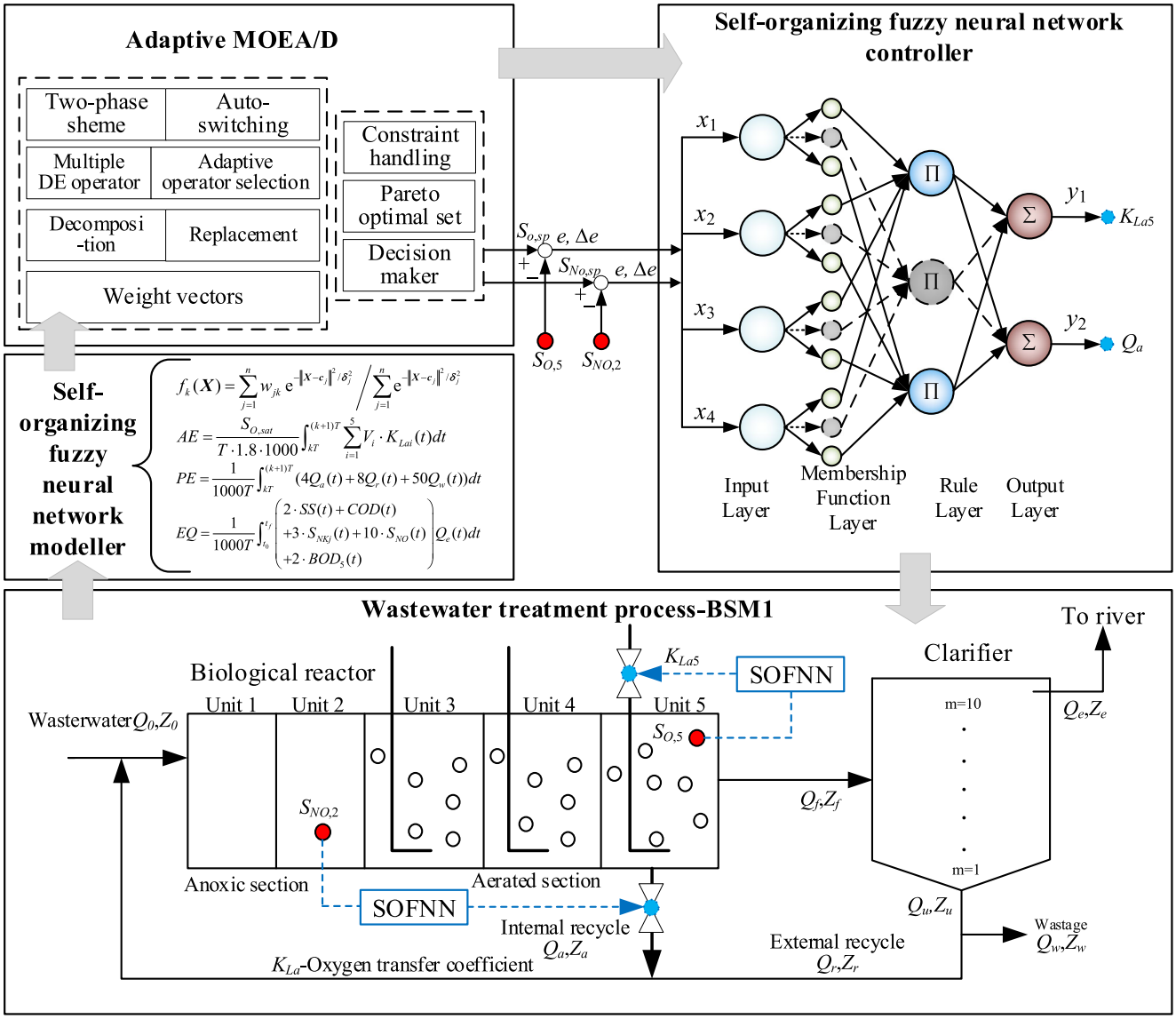

Fig. 6 Architecture of HMOOC for the WWTP

Step 1: The optimization objectives are established using the SOFNN-based prediction models with the process data in BSM1. The inputs of the models are the five decision variables and influent quality parameters. And the outputs of the models are the EC, EQ and effluent parameters.

$$f_k(\mathbf{x}) = \sum_{j=1}^{u} w_{jk}\varphi_j = \frac{\sum_{j=1}^{u}\left[w_{jk}\exp\left(-\frac{\|\mathbf{x}-\mathbf{c}_j\|^2}{\delta_j^2}\right)\right]}{\sum_{j=1}^{u}\exp\left(-\frac{\|X-\mathbf{c}_j\|^2}{\delta_j^2}\right)} \quad (27)$$

where $\mathbf{x}=[x_1x_2,...,x_n]^{\mathrm{T}}$ is the input vector of SOFNN, n is the number of input variables, $\mathbf{c}_j=[c_{1j}c_{2j},...,c_{nj}]$ and $\delta_j=[\delta_{1j}\delta_{2j},...,\delta_{nj}]$ are the center vector and width vector of the jth rule neuron, respectively φ_j is the normalized output of the jth rule neuron, w_{jk} is the weight coefficient between the jth rule neuron and the kth output neuron $j=1,2,...,uu$ is the number of rule neurons, and $k=1,2,...,ss$ is the number of output variables. In SOFNN, the network structure is dynamically adjusted by a self-organizing mechanism designed based on the singular value decomposition method. Meanwhile, the network parameters are optimized by an adaptive learning algorithm developed based on the improved Levenberg-Marquardt (LM) optimization method [43, 44].

Step 2: The optimization objectives are minimized by using the proposed AMOEA/D to obtain a set of Pareto optimal solutions.

Step 3: The intelligent decision system designed by the fuzzy membership function approach is utilized to select a preferred solution from the Pareto solution set. Then, the optimal set-points of $S_{O,5}$ and $S_{NO,2}$ in the current optimal cycle could be determined.

For the ith objective function f_i, the satisfaction of nondominated solution $f(\mathbf{x}_k)$ is defined as follows:

$$\mu_i^k = \begin{cases} 1, & f_i(\mathbf{x}_k) \le f_i^{\min} \\ \frac{f_i^{\max} - f_i(\mathbf{x}_k)}{f_i^{\max} - f_i^{\min}}, & f_i^{\min} < f_i(\mathbf{x}_k) < f_i^{\max} \\ 0, & f_i(\mathbf{x}_k) \ge f_i^{\max} \end{cases} \quad (28)$$

Table 1 Test instances F1 - F6

Instance	Description	Domain	Number of variables	Notes
F1	$f_1(x) = (1 + g(x))x_1$ $f_2(x) = (1 + g(x))(1 - \sqrt{x_1})^3$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^{n}(y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2:n} = x_i - \sin(0.5\pi x_i)$ POF: $f_2 = (1 - \sqrt{f_1})^3$ POS: $x_i = \sin(0.5\pi x_i), i = 2, \ldots, n$	$x_i \in [0,1]\ 1 \le i \le n$	20	Unimodal convex
F2	$f_1(x) = (1 + g(x))x_1$ $f_2(x) = (1 + g(x))\sqrt{1 - x_1^5}$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^{n}(y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2:n} = x_i - \sin(0.5\pi x_i)$ POF: $f_2 = \sqrt{1 - f_1^5}$ POS: $x_i = \sin(0.5\pi x_i), i = 2, \ldots, n$	$x_i \in [0,1]\ 1 \le i \le n$	20	Unimodal concave
F3	$f_1(x) = (1 + g(x))x_1$ $f_2(x) = \frac{1}{2}(1 + g(x))(1 - x_1^{0.1} + (1 - \sqrt{x_1})^2\cos^2(3\pi x_1))$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^{n}(y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2:n} = x_i - \sin(0.5\pi x_i)$ POF: $f_2 = \frac{1}{2}(1 - f_1^{0.1} + (1 - \sqrt{f_1})^2\cos^2(3\pi f_1))$ POS: $x_i = \sin(0.5\pi x_i), i = 2, \ldots, n$	$x_i \in [0,1]\ 1 \le i \le n$	20	Multimodal disconnected
F4	$f_1(x) = (1 + g(x))(x_1 + 0.05\sin(6\pi x_1))^2$ $f_2(x) = (1 + g(x))(1 - x_1 + 0.05\sin(6\pi x_1))^2$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^{n}(y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2:n} = x_i - \sin(0.5\pi x_i)$ POF: $f_1^{0.5} + f_2^{0.5} = 1 + 0.1\sin(3\pi(f_1^{0.5} - f_2^{0.5} + 1))$ POS: $x_i = \sin(0.5\pi x_i), i = 2, \ldots, n$	$x_i \in [0,1]\ 1 \le i \le n$	20	Multimodal mixed
F5	$f_1(x) = (1 + g(x))(x_1 + 0.05\sin(6\pi x_1))^{0.2}$ $f_2(x) = (1 + g(x))(1 - x_1 + 0.05\sin(6\pi x_1))^{10}$ $g(x) = 2\sin(0.5\pi x_1)(n - 1 + \sum_{i=2}^{n}(y_i^2 - \cos(2\pi y_i)))$ where $y_{i=2:n} = x_i - \sin(0.5\pi x_i)$ POF: $f_1^5 + f_2^{0.1} = 1 + 0.1\sin(3\pi(f_1^5 - f_2^{0.1} + 1))$ POS: $x_i = \sin(0.5\pi x_i), i = 2, \ldots, n$	$x_i \in [0,1]\ 1 \le i \le n$	20	Multimodal mixed
F6	$f_1(x) = ((1 + g(x))\cos(0.5\pi x_1)\cos(0.5\pi x_2))^4$ $f_2(x) = ((1 + g(x))\cos(0.5\pi x_1)\sin(0.5\pi x_2))^4$ $f_3(x) = ((1 + g(x))\sin(0.5\pi x_1))^2$ $g(x) = \sum_{i=3}^{n}(x_i - 0.5)^2$ POF: $\sqrt{f_1} + \sqrt{f_2} + f_3 = 1$ POS: $x_i = 0.5, i = 3, \ldots, n$	$x_i \in [0,1]\ 1 \le i \le n$	20	Unimodal convex

where $f_i^{\max}$ and f_i^{min} are the maximum and minimum of the ith objective function f_i respectively. The normalized satisfaction of $f(\mathbf{x}_k)$ is as follows:

$$\mu^k = \frac{\sum_{i=1}^{m} \mu_i^k}{\sum_{k=1}^{|S|} \sum_{i=1}^{m} \mu_i^k} \tag{29}$$

where m is the number of objectives and $|S|$is the number of elements in POS obtained. In this study, the solution with the maximum value of μ^k is selected by the intelligent decision system as the preferred solution.

Step 4: The optimal set-points of $S_{O,5}$ and $S_{NO,2}$ is tracked by the intelligent multivariable controller, which is designed using self-organizing fuzzy neural network [45]. The SOFNN controller displaying high steady-state accuracy and strong self-adaptability under complex conditions, can meet the needs of the bottom control loops. If the continuous 14-days data simulation is completed in BSM1, then stop. Otherwise, go to **Step 1** for the next optimal cycle.

5 Benchmark problems testing

5.1 Test instances

To verify the validity of the proposed AMOEA/D algorithm, twelve unconstrained MOPs with complex POFs are used for testing. Table 1 gives the detailed description of the first six test instances (F1-F6) which were designed in [34] and Table 2 presents the detailed description of the remaining test instances (F7-F12) which were respectively adopted in [35] and [36].

5.2 Parameter settings

In this paper, the proposed AMOEA/D is compared with four state-of-the-art decomposition-based MOEAs, including MOEA/D-DE, MOEA/D-APS, MOEA/D-TPN, and iMOEA/D. In MOEA/D-DE, MOEA/D-TPN, and iMOEA/D, the Tchebycheff approach explained in (13) is adopted as the decomposition approach. MOEA/D-APS adopts the PBI-based decomposition method with adaptive penalty scheme. The parameters of MOEA/D-DE, MOEA/D-APS, MOEA/D-TPN, and iMOEA/D are set according to their corresponding references [34, 35, 37, 38], respectively. The detailed parameter settings of the proposed AMOEA/D are summarized as follows.

1) *Control parameters in polynomial mutation:* $p_m = 1/n$ and $\eta = 20$.
2) *Neighborhood size:* $T = 20$.
3) *Probability to select in the neighborhood:* $\delta = 0.9$.
4) *Control parameter in the replacement operator:* $n_r = 2$.
5) *Population size:* $N = 100$ for two-objective test instances, 300 for the three-objective ones.
6) Maximum number of iterations: *maxIteration* = 300 for F1-F5, 600 for F6-F12.
7) *Number of runs*: Each algorithm is run 30 times independently on each test instance.

5.3 Performance metric

In our empirical studies, we consider the following two widely used performance metrics [34].

1) Inverted Generational Distance (IGD)

Let S^* be a set of points uniformly sampled from the true POF, and S be the set of approximated solutions obtained by an MOEA, the IGD indicator measures the gap between S^* and S, which is calculated as follows:

$$IGD(S^*, S) = \frac{\sum_{\mathbf{x} \in S^*} d(\mathbf{x}, S)}{|S^*|} \tag{30}$$

where $d(\mathbf{x}, S)$ is the Euclidean distance between the solution $\mathbf{x}$ and its nearest neighbor in S, and $|S^*|$ is the cardinality of S^*. If the number of points in S^* is big enough, the IGD indicator can measure the convergence and diversity of the approximated POF obtained by an MOEA at the same time. The smaller the IGD value, the better the quality of the approximated POF. In the experimental studies, 500 uniformly distributed points are sampled from the true POF for bi-objective test instances, and 1000 for three-objective ones, respectively.

2) Hypervolume (HV)

Let $\mathbf{z}^r = (z_1^r, z_2^r, ..., z_m^r)^T$ be a reference point in the objective space that is dominated by all points on the true POF, and S be the set of approximated solutions obtained by an MOEA, the HV indicator measures the size of objective space dominated by the solution in S and bounded by $\mathbf{z}^r$

$$HV(S) = VOL(\bigcup_{\mathbf{x} \in S} [f_1(\mathbf{x}), z_1^r] \times ... [f_m(\mathbf{x}), z_m^r]) \tag{31}$$

where VOL represents the Lebesgue measure. The bigger the HV value, the better the quality of the approximated POF. In the experimental studies, $\mathbf{z}^r$ is set to $(1.1, 1.1)^T$ for F1-F5 $(1.1, 1.1, 1.1)^T$ for F6, $(1.1, 11)^T$ for F7, $(1.1, 1.1)^T$ for F8-F9, $(5.5, 5.5, 5.5)^T$ for F10, and $(1.1, 1.1, 1.1)^T$ for F11-F12 when computing HV for the nondominated sets obtained by all the algorithms.

Table 2 Test instances F7 - F15

Instance	Description	Domain	Number of variables	Notes
F7	$f_1 = 1 - \cos(0.5\pi x_1) + \frac{2}{\lvert J_1\rvert}\sum_{j\in J_1}\lvert y_j\rvert^{0.7}$ $f_2 = 10 - 10\sin(0.5\pi x_1) + \frac{2}{\lvert J_2\rvert}\sum_{j\in J_2}\lvert y_j\rvert^{0.7}$ $y_j = x_j - 0.9\sin\left(\frac{j\pi}{n}\right), j \in 2, ..., n.$ where $J_1 = \{j \mid j\ is\ odd\ and\ 2 \le j \le n\}\ and\ J_2 = \{j \mid j\ is\ even\ and\ 2 \le j \le n\}$	$x_1 \in [0,1]\ x_i \in [-1,1]\ 2 \le i \le n$	30	Unimodal convex
F8	$f_1 = x_1 + \frac{2}{\lvert J_1\rvert}\sum_{j\in J_1}\left\lvert y_j + \frac{\sin(\pi y_j)}{\pi}\right\rvert$ $f_2 = \begin{cases} 1 - 19x_1 + \frac{2}{\lvert J_2\rvert}\sum_{j\in J_2}\left\lvert y_j + \frac{\sin(\pi y_j)}{\pi}\right\rvert & if\ x_1 \le 0.005 \\ \frac{1}{19} - \frac{x}{19} + \frac{2}{\lvert J_2\rvert}\sum_{j\in J_2}\left\lvert y_j + \frac{\sin(\pi y_j)}{\pi}\right\rvert & otherwise \end{cases}$ $y_j = x_j - 0.9\sin\left(\frac{j\pi}{n}\right),\ j \in 2, ..., n.$ where J_1 and J_2 are the same with as those of F7.	$x_1 \in [0,1]\ x_i \in [-1,1]\ 2 \le i \le n$	30	Unimodal convex
F9	$f_1 = x_1 + \frac{2}{\lvert J_1\rvert}\sum_{j\in J_1}(1 - e^{-\lvert y_j\rvert})$ $f_2 = \begin{cases} 1 - 8x_1^4 + \frac{2}{\lvert J_2\rvert}\sum_{j\in J_2}(1 - e^{-\lvert y_j\rvert}) & if\ x_1 \le 0.5 \\ 8(1 - x_1)^4 + \frac{2}{\lvert J_2\rvert}\sum_{j\in J_2}(1 - e^{-\lvert y_j\rvert}) & otherwise \end{cases}$ $y_j = x_j - 0.9\sin\left(\frac{j\pi}{n}\right),\ j \in 2, ..., n.$ where J_1 and J_2 are the same with as those of F7.	$x_1 \in [0,1]\ x_i \in [-1,1]\ 2 \le i \le n$	30	Unimodal mixed
F10	$f_1 = (1 - g(x))(\frac{x_1}{\sqrt{x_2x_3}})$ $f_2 = (1 - g(x))(\frac{x_2}{\sqrt{x_1x_3}})$ $f_3 = (1 - g(x))(\frac{x_3}{\sqrt{x_1x_2}})$ $g(x) = \sum_{i=4}^{n}(x_i - 2)^2$	$x_i \in [1,4]\ 1 \le i \le n$	30	Multimodal convex separable
F11	$f_1 = (1 + g(x))((1 - x_1)x_2)$ $f_2 = (1 + g(x))(x_1(1 - x_2))$ $f_3 = (1 - g(x))(1 - x_1 - x_2 + 2x_1x_2)^6$ $g(x) = \sum_{i=3}^{n}(x_i - 0.5)^2$	$x_i \in [0,1]\ 1 \le i \le n$	30	Unimodal convex non-separable
F12	$f_1 = \cos^4(0.5\pi x_1)\cos^4(0.5\pi x_2)$ $f_2 = \cos^4(0.5\pi x_1)\sin^4(0.5\pi x_2)$ $f_3 = \left(\frac{1}{1+\cos^2(0.5\pi x_1)}\right)^{\frac{1}{1+g(x)}}$ $g(x) = \frac{1}{10}\sum_{i=3}^{n}(1 + x_i^2 - \cos(2\pi x_i))$	$x_i \in [0,1]\ 1 \le i \le n$	3	Multimodal convex separable

Table 3 IGD comparison results

Prob.		IGD				
		MOEA/D-DE [38]	MOEA/D-APS [34]	MOEA/D-TPN [35]	iMOEA/D [37]	AMOEA/D
F1	Best	0.0228	0.0615	0.0054	0.0052	**0.0045**
	mean	0.0232 [a]	0.0634 [a]	0.0058 [a]	0.0060 [a]	**0.0045**
	worst	0.0235	0.0654	0.0062	0.0067	**0.0046**
F2	Best	0.0047	0.0073	0.0061	0.0059	**0.0057**
	mean	0.6280 [a]	0.3818 [a]	0.1783 [a]	0.0922 [a]	**0.0613**
	worst	**0.7838**	**0.7838**	**0.7838**	**0.7838**	**0.7838**
F3	Best	0.0190	0.1654	0.0049	0.0046	**0.0036**
	mean	0.0330 [a]	0.3160 [a]	0.0052	0.0055 [a]	**0.0038**
	worst	0.3865	0.3942	0.0058	0.0059	**0.0039**
F4	Best	0.0120	0.0718	0.0047	0.0047	**0.0045**
	mean	0.0614 [a]	0.1179 [a]	0.0051 [a]	0.0049	**0.0046**
	worst	0.7525	0.7525	0.0055	0.0050	**0.0048**
F5	Best	0.0069	0.095	0.0060	**0.0053**	0.0055
	mean	**0.0075**	0.0156 [a]	0.122 [a]	0.0108 [a]	0.0080
	worst	**0.0088**	0.0148	0.232	0.0195	0.0153
F6	Best	0.698	**0.0277**	0.0311	–	0.0325
	mean	0.0702 [a]	**0.0286** [b]	0.0335 [a]	–	0.0344
	worst	0.0705	**0.0291**	0.354	–	0.0417

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-AAP and each of MOEA/D-DE, MOEA/D-STM, MOEA/D-ACD and MOEA/D-APS. [a] and [b] denote the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-AAP, respectively

–denotes that the iMOEA/D should not be used to solve three-objective optimization problem

Table 4 HV comparison results

Prob.		HV				
		MOEA/D-DE [38]	MOEA/D-APS [34]	MOEA/D-TPN [35]	iMOEA/D [37]	AMOEA/D
F1	Best	1.1050	1.1007	1.1069	1.1070	**1.1072**
	mean	1.1049 [a]	1.1003 [a]	1.1068	1.1068	**1.1070**
	worst	1.1049	1.0997	1.1065	1.1066	**1.1069**
F2	Best	**0.3109**	0.3046	0.3046	0.3059	**0.3077**
	mean	0.1502 [a]	0.1934 [a]	0.2099 [a]	0.2204 [a]	**0.2879**
	worst	**0.1100**	**0.1100**	**0.1100**	**0.1100**	**0.1100**
F3	Best	1.1263	1.1157	1.1251	1.1265	**1.1280**
	mean	1.1241 [a]	1.0944 [a]	1.1250 [a]	1.1262 [a]	**1.1280**
	worst	1.0853	1.0834	1.1248	1.1260	**1.1279**
F4	Best	1.0351	1.0306	1.0360	1.0359	**1.0362**
	mean	0.9733 [a]	0.9690 [a]	1.0359	1.0358	**1.0360**
	worst	0.1100	.1100	1.0358	1.0358	**1.0360**
F5	Best	0.5322	0.5329	0.5333	0.5332	**0.5337**
	mean	0.5321 [a]	0.5302 [a]	0.5330	0.5329	**0.5336**
	worst	0.5321	0.5285	0.5325	0.5327	**0.5334**
F6	Best	1.2660	**1.2854**	1.2818	–	1.2802
	mean	1.2655 [a]	**1.2846** [b]	1.2790	–	1.2788
	worst	1.2652	**1.2836**	1.2766	–	1.2749

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-AAP and each of MOEA/D-DE, MOEA/D-STM, MOEA/D-ACD and MOEA/D-APS. [a] and [b] denote the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-AAP, respectively

–denotes that the iMOEA/D should not be used to solve three-objective optimization problem

5.4 Experimental results

5.4.1 Comparisons on F1-F6

Tables 3 and 4 give the best, average, and worst values of IGD and HV for MOEA/D-DE, MOEA/D-APS MOEA/D-TPN, iMOEA/D, and AMOEA/D on F1-F6 test instances respectively, in which the bold means the corresponding algorithm achieves the best results on the test instance. The differences between the approximations are assessed by the Wilcoxon rank-sum test at the 0.05 significance level. Signs of ^{a} and ^{b} in the superscript form on mean values indicate the significance of the proposed algorithms. From Table 3, the proposed AMOEA/D can effectively tackle the complex MOPs. For F1, F3-F5, compared with MOEA/D-DE and MOEA/D-APS, the IGD values of AMOEA/D decrease significantly, indicating that the TP optimization with auto-switching scheme is more beneficial to find boundary solutions and to improve the population diversity. Compared with MOEA/D-TPN and iMOEA/D, the performance of AMOEA/D has a slight improvement since the adopted adaptive DE strategy and auto-switching scheme can enhance the quality of POS on F1-F4. For F2, it is easy for MOEA/D-DE and MOEA/D-APS to fall into a local POF, resulting in the high mean IGD values, while AMOEA/D has a stable performance and can approximate the whole POF in most cases. For the three-objective optimization problem F6, compared with MOEA/D-DE, the IGD values of MOEA/D-APS, MOEA/D-TPN, and AMOEA/D algorithms decrease to a certain extent. From Table 4, the similar conclusions can be drawn.

Figure 7 shows the approximated POFs for MOEA/D-DE, MOEA/D-APS, and AMOEA/D when they obtain the lowest IGD values on F1 to F6. From Fig. 7, the solutions obtained by MOEA/D-DE and MOEA/D-APS might converge to the middle region of the POFs, resulting in the decrease of population diversity. The proposed AMOEA/D can obtain a better solution distribution along the actual POF, especially for F1, F3, and F4 test instances with convex POFs. It is worth noting that when the

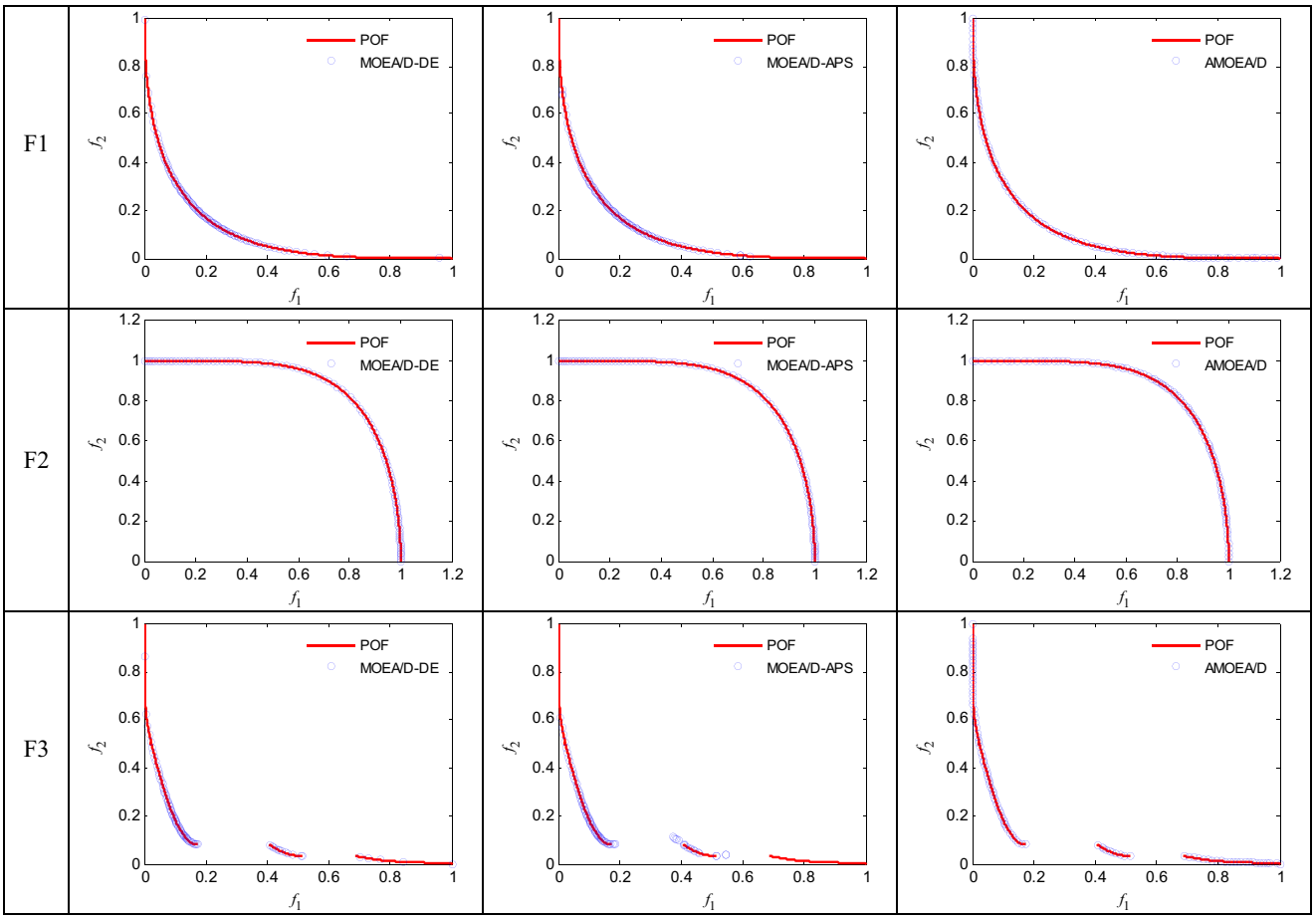

Fig. 7 The obtained approximated POF on F1-F6 when the minimum IGD is obtained during 30 times' running

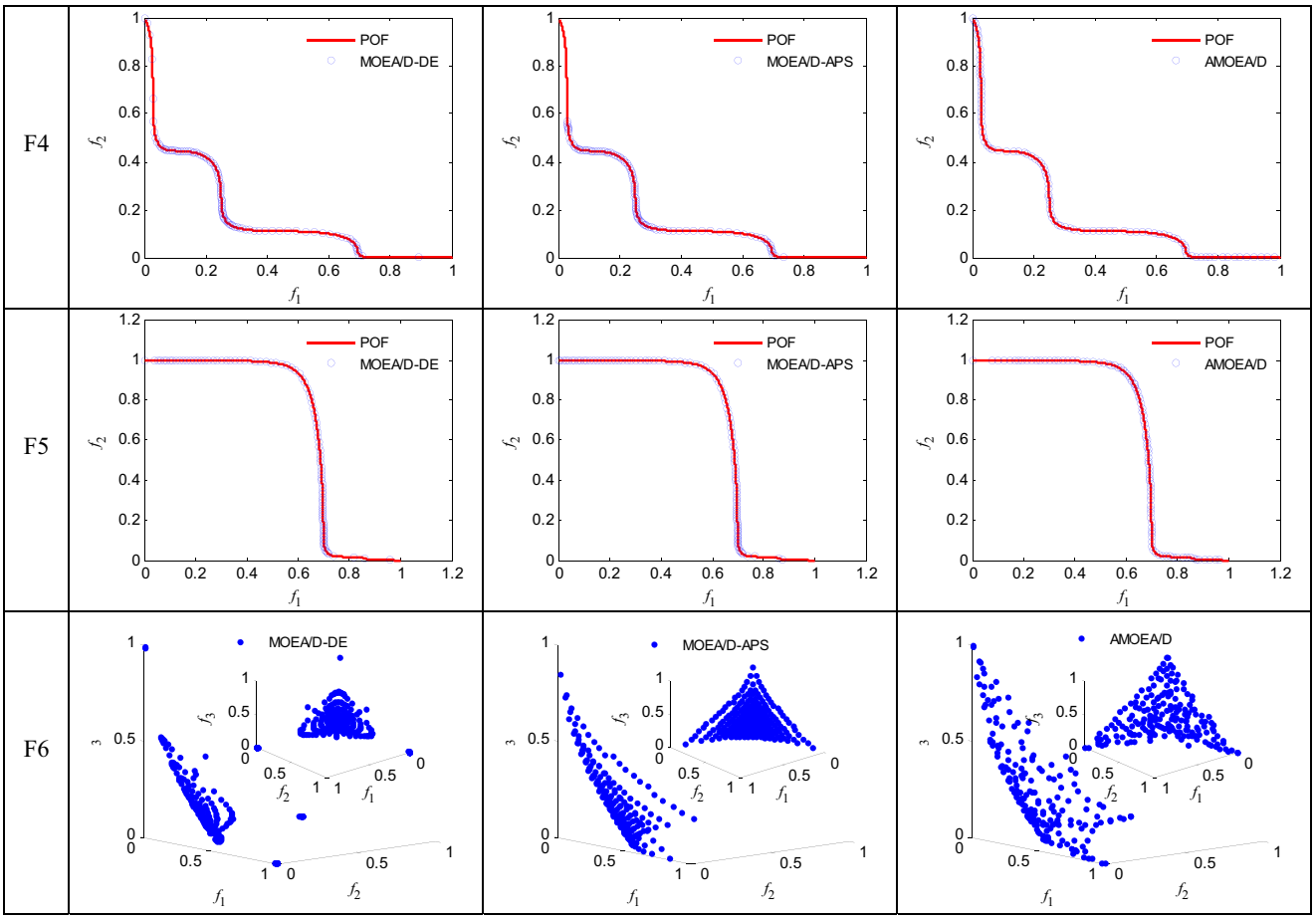

Fig. 7 (continued)

minimum IGD value is obtained, MOEA/D-DE, MOEA/D-APS, and AMOEA/D achieve the same performance on F2 with a concave POF.

In order to more clearly illustrate the proposed strategies, Fig. 8 plots the evolution curves of the IGD metric value versus the number of generations for each algorithm on F1 and F4 instances when the minimum IGD is obtained during the 30 times' running. It can be observed that the proposed AMOEA/D can offer a significant improvement on the IGD metric when the TP optimization with auto-switching scheme is activated.

Figure 9 presents the boxplots of the IGD metric obtained by the five algorithms from 30 independent runs on each instance. These results clearly indicate that the AMOEA/D is the best on F1-F5. And, more remarkably, for F2, the solutions obtained by MOEA/D-DE and MOEA/D-APS might converge to the peak point of the POF, while the solutions obtained by the proposed AMOEA/D can cover

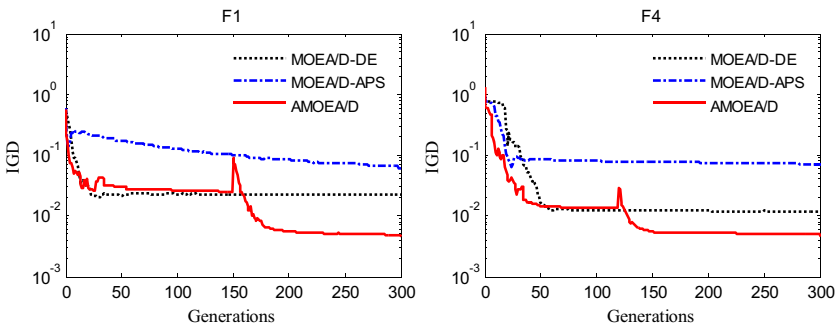

Fig. 8 The IGD evolution curves of F1 and F4 when the minimum IGD value is obtained during the 30 times' running

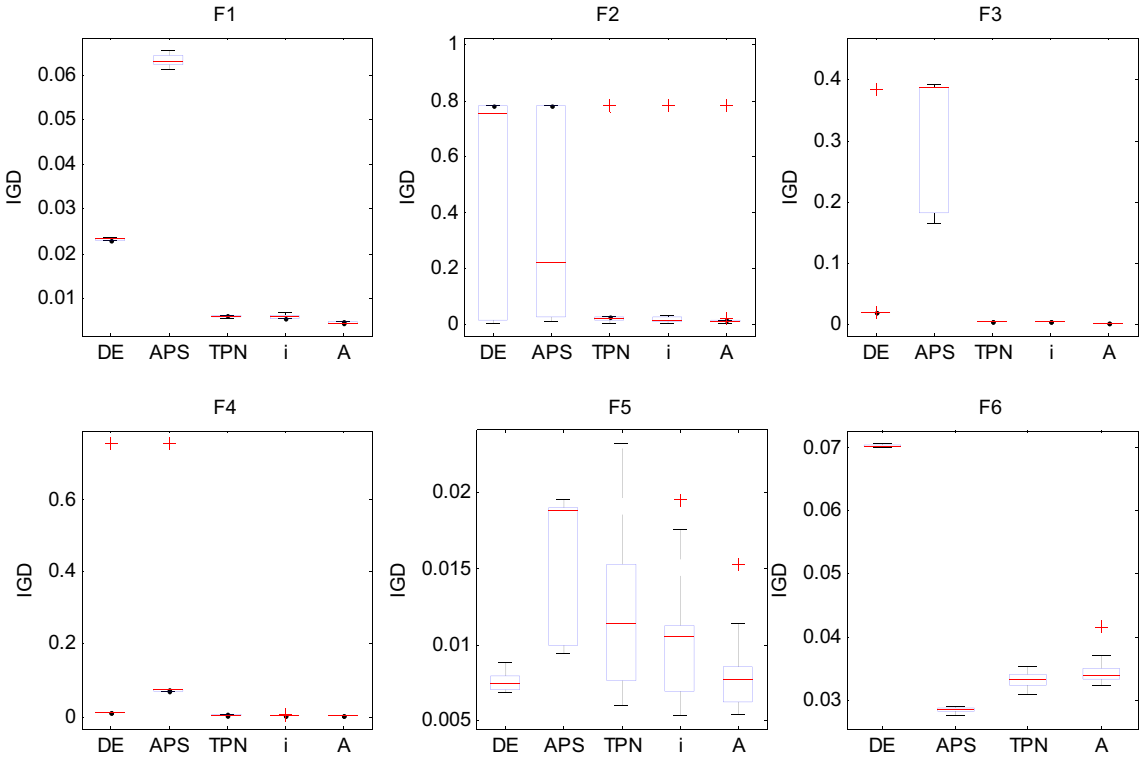

Fig. 9 Boxplots for the comparison of AMOEA/D with MOEA/D-DE, MOEA/D-APS, MOEA/D-TPN, and iMOEA/D on IGD. DE, APS, TPN, i, and A in the horizontal axis stand for MOEA/D-DE, MOEA/D-APS, MOEA/D-TPN, iMOEA/D, and AMOEA/D, respectively

Table 5 IGD comparison results

Prob.		IGD				
		MOEA/D-DE [38]	MOEA/D-APS [34]	MOEA/D-TPN [35]	iMOEA/D [37]	AMOEA/D
F7	Best	0.0256	0.1431	0.0070	0.0096	**0.0063**
	mean	0.0276 ^{a}	0.1597 ^{a}	0.0088 ^{a}	0.0103 ^{a}	**0.0068**
	worst	0.0283	0.1796	0.0098	0.0112	**0.0081**
F8	Best	0.0175	0.2592	0.0061	0.0062	**0.0051**
	mean	0.0182 ^{a}	0.2773 ^{a}	0.0075 ^{a}	0.0079 ^{a}	**0.0052**
	worst	0.0201	0.3113	0.0092	0.0112	**0.0055**
F9	Best	0.0089	0.0195	**0.0054**	0.0060	0.0055
	mean	0.0114 ^{a}	0.0225 ^{a}	0.0065	0.0073 ^{a}	**0.0063**
	worst	0.0141	0.0265	0.0088	0.0108	**0.0082**
F10	Best	0.0278	0.0245	0.0231	–	**0.0211**
	mean	0.0279 ^{a}	0.0260 ^{a}	0.0249 ^{a}	–	**0.0212**
	worst	0.0280	0.0273	0.0262	–	**0.0212**
F11	Best	0.0503	0.0436	**0.0149**	–	0.0150
	mean	0.0506 ^{a}	0.0461 ^{a}	0.0165 ^{a}	–	**0.0155**
	worst	0.0510	0.0476	0.0186	–	**0.0165**
F12	Best	0.0724	1.2127	**0.0234**	–	0.0238
	mean	0.0735 ^{a}	1.2217 ^{a}	**0.0395** ^{b}	–	0.0408
	worst	0.0759	1.2252	0.0956	–	**0.0649**

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-AAP and each of MOEA/D-DE, MOEA/D-STM, MOEA/D-ACD and MOEA/D-APS. ^{a} and ^{b} denote the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-AAP, respectively

−denotes that the iMOEA/D can not be used to solve three-objective optimization problem

Bold means the corresponding algorithm achieves the best results on the test instance

Table 6 HV comparison results

Prob.		HV				
		MOEA/D-DE [38]	MOEA/D-APS [34]	MOEA/D-TPN [35]	iMOEA/D [37]	AMOEA/D
F7	Best	9.6906	8.7655	9.8479	9.7709	9.8468
	mean	9.6769 ^{a}	8.5676 ^{a}	9.8074 ^{a}	9.7274 ^{a}	**9.8378**
	worst	9.6665	8.3319	9.7654	9.6854	**9.8176**
F8	Best	1.1559	1.1264	1.1572	1.1569	**1.1575**
	mean	1.1557 ^{a}	1.1244 ^{a}	1.1570 ^{a}	1.1567 ^{a}	**1.1572**
	worst	1.1550	1.1206	1.1562	1.1560	**1.1568**
F9	Best	**0.7019**	0.6889	0.7013	0.7006	0.7014
	mean	**0.7009** ^{b}	0.6855 ^{a}	0.7007	0.7002 ^{a}	0.7007
	worst	**0.6999**	0.6819	0.6988	0.6988	0.6994
F10	Best	65.006	65.223	65.224	–	**65.227**
	mean	64.989 ^{a}	64.989 ^{a}	64.993 ^{a}	–	**65.212**
	worst	64.970	64.329	64.810	–	**65.199**
F11	Best	1.2955	1.3038	1.3061	–	**1.3064**
	mean	1.2954 ^{a}	1.3029 ^{a}	1.3057 ^{a}	–	**1.3060**
	worst	1.2952	1.3023	1.3039	–	**1.3050**
F12	Best	0.7045		0.7164	–	**0.7185**
	mean	0.7043 ^{a}	^{a}	**0.7130** ^{a}	–	0.7128
	worst	0.7039		0.793	–	**0.7120**

Wilcoxon's rank sum test at a 0.05 significance level is performed between MOEA/D-AAP and each of MOEA/D-DE, MOEA/D-STM, MOEA/D-ACD and MOEA/D-APS. † and ‡ denote the performance of the corresponding algorithm is significantly worse than or better than that of MOEA/D-AAP, respectively

–denotes that the iMOEA/D can not be used to solve three-objective optimization problem

Bold means the corresponding algorithm achieves the best results on the test instance

the entire POF in most cases. Thus, we can conclude that the proposed AMOEA/D has successfully improved the algorithm performance on MOPs with complicated POFs.

5.4.2 Comparisons on F7-F12

To verify the robustness of AMOEA/D on MOPs with different shapes of POFs, AMOEA/D is compared with MOEA/D-DE, MOEA/D-APS, MOEA/D-TPN, and iMOEA/D on F7-F12 test instances in this section. Tables 5 and 6 show the performance of the five compared algorithms in terms of IGD and HV on F7-F12, in which the bold means the corresponding algorithm achieves the best results on the test instance. It can be observed that the proposed AMOEA/D performs significantly better than MOEA/D-DE and MOEA/D-APS on all test instances in terms of IGD. From Table 5, AMOEA/D has the best performance on F7, F8, and F11, and MOEA/D-TPN has the best performance on F12. The performance of AMOEA/D is very similar to that of MOEA/D-TPN on F9. Similar performance can be observed on the comparisons of five algorithms in terms of HV, where AMOEA/D performs better than the other

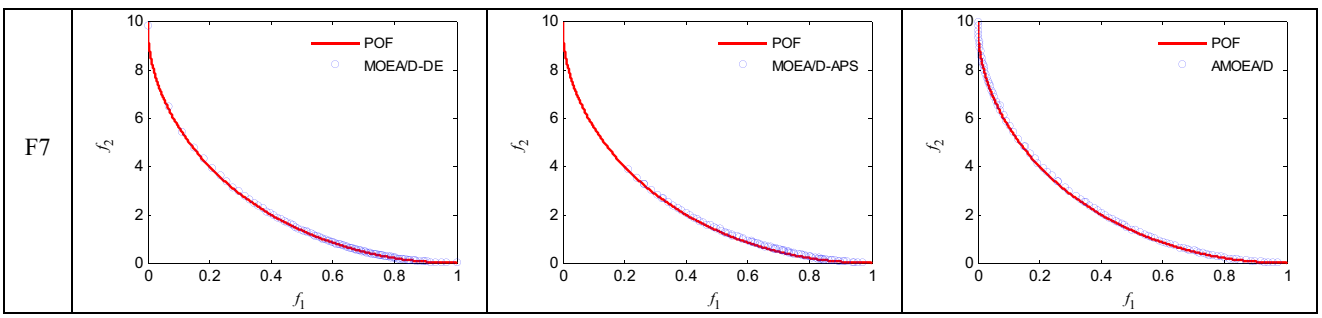

Fig. 10 The obtained approximated POF on F7-F15 when the minimum IGD is obtained during 30 times' running

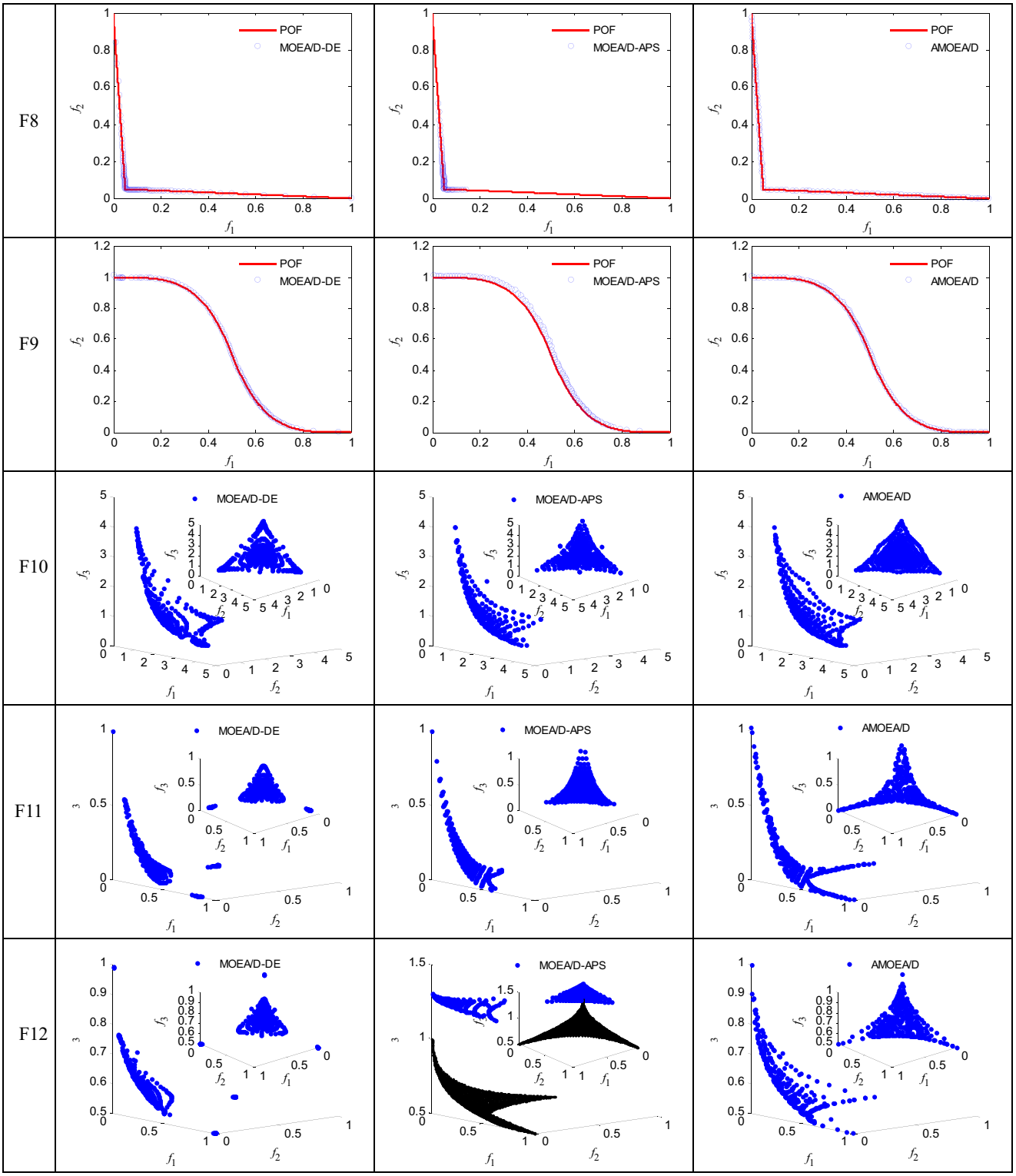

Fig. 10 (continued)

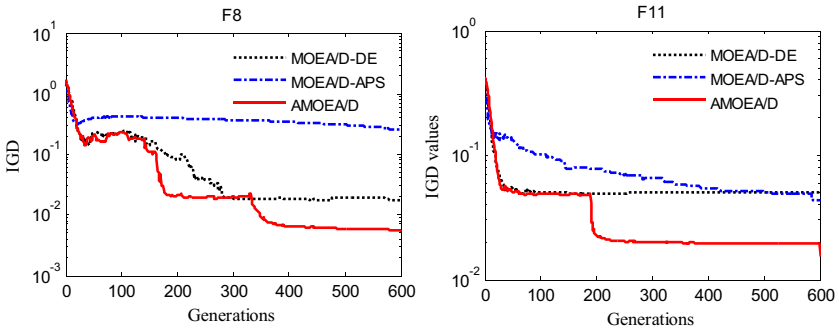

Fig. 11 The IGD evolution curves of F10 and F12 when the minimum IGD value is obtained during the 30 times' running

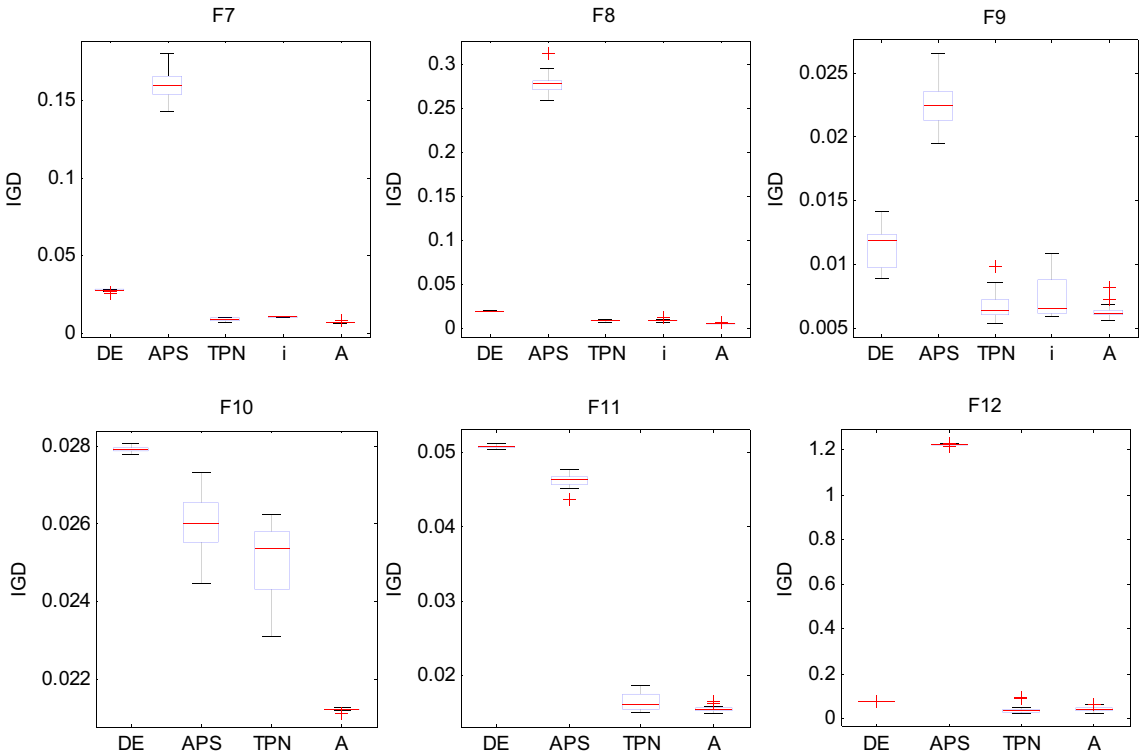

Fig. 12 Boxplots for the comparison of AMOEA/D with MOEA/D-DE, MOEA/D-APS, MOEA/D-TPN, and iMOEA/D on IGD. DE, APS, TPN, i, and A in the horizontal axis stand for MOEA/D-DE, MOEA/D-APS, MOEA/D-TPN, iMOEA/D, and AMOEA/D, respectively

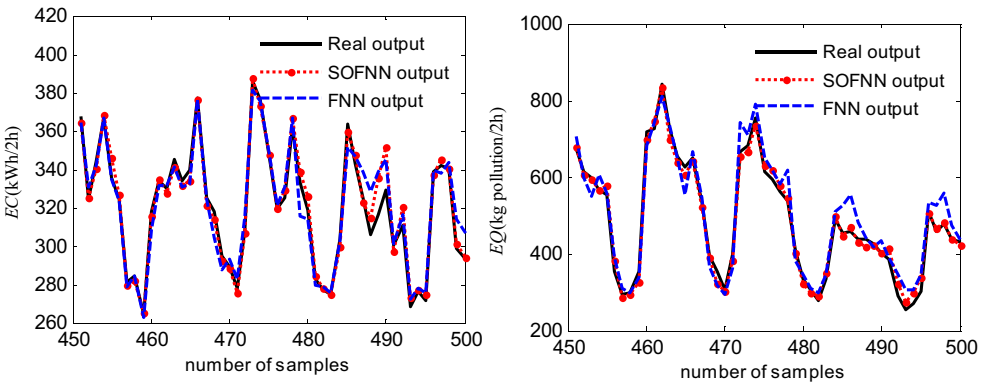

Fig. 13 The predicting results for EC and EQ

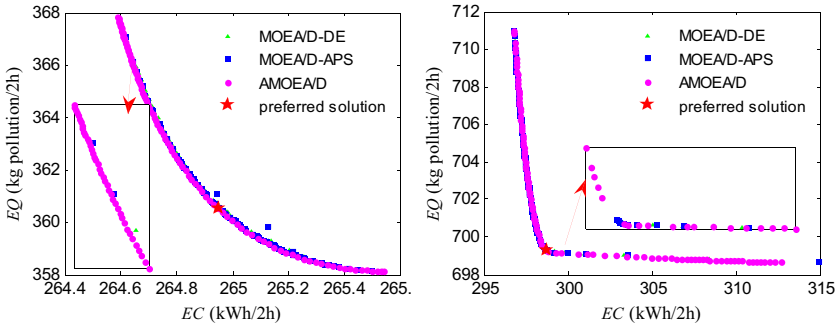

Fig. 14 Approximated POFs obtained by MOEA/D, MOEA/D-APS, and AMOEA/D

compared algorithms on four out of six test instances. This further validates the efficiency of the proposed AMOEA/D algorithm in solving complex problems tested in this paper.

Figure 10 shows the approximated POFs achieved by MOEA/D-DE, MOEA/D-APS, and AMOEA/D with the best IGD value on each test instance. It can be seen from Fig. 10 that F7, F8 and F9 have disparately scaled objectives, extremely convex POF and mixed POF, respectively. Furthermore, F10, F11 and F12 are three three-objective MOPs with complicated POFs, in which the boundary parts of their POFs are more difficult to approximate than other parts. Thus, it is desirable to obtain a more comprehensive evaluation on the performance of the compared algorithms by using these test problems with different characteristics. It can be observed that the solutions obtained by MOEA/D-DE and MOEA/D-APS might converge to the part region of the POFs on F7-F12 expect for F9, resulting in the decrease of population diversity. The proposed AMOEA/D can obtain a better solution distribution along the actual POF. It is worth noting that MOEA/D-APS with a smaller initial penalty factor is unable to find the actual POF on F12 due to the existence of many local POFs.

To compare the convergence speed of the compared algorithms, Fig. 11 plots the evolution curves of the IGD metric value versus the number of generations for MOEA/D-DE, MOEA/D-APS, and AMMOEA/D on F8 and F11 instances when the minimum IGD is obtained during the 30 times' running. It can be observed that AMOEA/D converges faster than MOEA/D-DE and MOEA/D-APS, since they are easy to trap into local POFs. Meanwhile, a significant improvement of the IGD values can be observed when the second phase is automatically activated.

Figure 12 shows the boxplots of the IGD metric obtained by the five algorithms from 30 independent runs on each instance. It can be observed that the performance of AMOEA/D is superior to MOEA/D-DE and MOEA/D-APS on all test problems. Besides, it can be seen from the boxplots of F7-F9 that the beards of MOEA/D-DE and MOEA/D-APS are longer than that of AMOEA/D. It means that MOEA/D-DE and MOEA/D-APS are still easy to trap into local POFs. Compared with MOEA/D-TPN, AMOEA/D obviously performs better on F7, F9, F10 and F11. Thus, we can conclude that the proposed AMOEA/D outperforms all the compared algorithms in most test instances. More remarkably, AMOEA/D is very robust with the shapes of POFs and can effectively solve MOPs with complicated POFs (e.g., discontinuous POFs or POFs with a sharp peak and long tail). It is interesting that MOEA/D-APS performs poorly on F7, F8 and F12 since it adopts PBI-based decomposition approach which is easily converged to the central region on MOPs with extremely convex POFs.

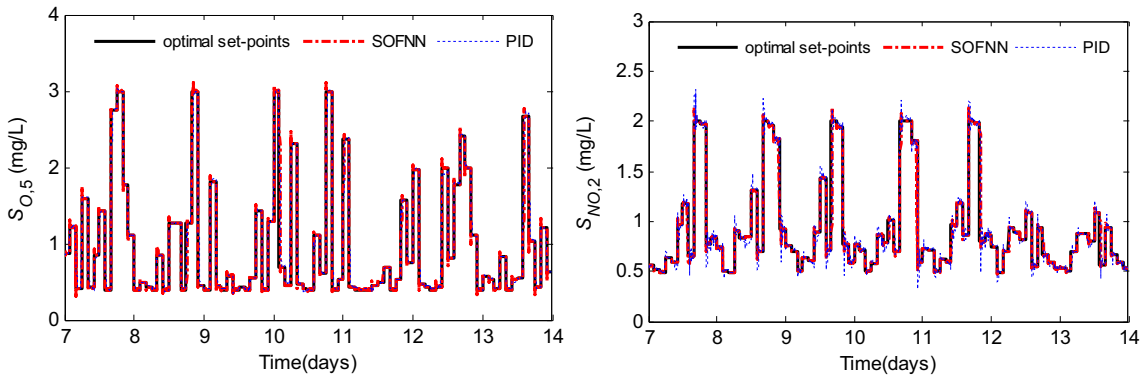

Fig. 15 Optimized set-points and online tracking control performance of S_O and S_{NO} in dry weather

6 Multiobjective optimal control in the WWTP

In this section, the proposed AMOEA/D is utilized to implement the multiobjective optimal control in the WWTP. The BSM1 is introduced to evaluate the effectiveness of the AMOEA/D-based HMOOC strategy. The sampling period of the SOFNN controller was set to 45 seconds and the optimal cycle was set to 2 hours. For AMOEA/D, the population size N was set to 100 and the maximum number of iteration was set to 300.

6.1 Modeling results

First, 500 data samples are generated with the usage of BSM1. Then, the SOFNN prediction models are adopted to establish the objective functions of EC, EQ and effluent parameters. The predicting results of EC and EQ with the usage of SOFNN and FNN are shown in Fig. 13. It can be observed that the prediction accuracy of SOFNN is higher than that of FNN. Obviously, the SOFNN-based approximator can provide more accurate objective functions for the HMOOC strategy. Specifically, for EC modeling, the testing RMSE of SOFNN and FNN are 5.49 and 7.87, respectively. For EQ modeling, the testing RMSE of SOFNN and FNN are 16.44 and 36.83, respectively.

6.2 Optimization results

The data file of dry weather is utilized to test the performance of different algorithms, and the approximated POFs obtained by MOEA/D-DE, MOEA/D-APS and AMOEA/D in some two optimal cycles are shown in Fig. 14. It can be seen that the convergence and diversity of solutions obtained by AMOEA/D is better than that of MOEA/D-DE and MOEA/D-APS. Especially, more boundary solutions can be found by AMOEA/D, which indicates that TP optimization with auto-switching scheme and adaptive DE strategy can enhance the search performance of MOEA/D. While using the intelligent decision system to select a preferred solution, the EC of AMOEA/D is lower than that of MOEA/D-DE and MOEA/D-APS.

6.3 Optimal control results

The variation of the optimized set-points and the tracking effect of the SOFNN controller in dry weather are presented in Fig. 15. It can be seen that $S_{O,5}$ and $S_{NO,2}$ can be dynamically adjusted with the change of influent condition and component concentration, achieving the dynamic balance between EC and EQ. Furthermore, the control effect of the SOFNN controller is better than that of the

Table 7 Comparison of EC and EQ with different optimal control strategies in dry weather

Weather	Method	S_{NH}	N_{tot}	BOD_5	COD	TSS	AE	PE	EC		EQ
		(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(kWh/d)	(kWh/d)	(kWh/d)	Up/Down	(kg poll./d)
Dry	Limits	4	18	10	100	30	–	–	–	–	–
	Influent	30.14	51.47	70.57	167.3	198.57	–	–	–	–	–
	PID	2.30	16.99	2.67	47.47	12.59	3676.28	231.33	3907.61	–	6096.63
	SOOC [24]	3.05*	16.31*	–	–	–	3592.87*	226.82*	3819.69*	↓ 2.25%*	6315.84*
	DDAOC [16]	2.73*	15.84*	3.11*	44.89*	–	–	–	3700.51*	↓ 5.30%*	–
	Hopfield [17]	3.24*	14.92*	2.69*	47.55*	12.62*	3435.11*	265.40*	3700.51*	↓ 5.30%*	–
	RTO-NMPC [18]	–	–	–	–	–	3559.00*	120.01*	3679.01*	↓ 5.85%*	6870.7*
	NSGAII [25]	3.19*	17.33*	2.69*	47.55*	12.62*	3517.80*	188.96*	3706.76*	↓ 5.14%*	6466.75*
	MOPSO [27]	317*	15.72*	2.88*	47.60*	1347*	3493.46	200.01*	3693.47	↓ 5.48%*	6322.82
	MOEA/D-DE [38]	2.73	15.22	2.68	47.54	12.58	3466.98	199.52	3666.50	↓ 6.17%	6318.04
	MOEA/D-APS [34]	2.73	15.33	2.69	47.56	12.58	3485.22	206.77	3691.99	↓ 5.52%	6328.18
	MOEA/D-TPN [35]	2.76	15.36	2.68	47.54	12.58	3431.03	234.69	366572	↓ 6.19%	6332.32
	iMOEA/D [37]	2.76	15.28	2.68	47.56	12.58	3477.40	205.52	368292	↓ 5.75%	6330.55
	AMOEA/D	2.75	15.36	2.68	47.54	12.58	3384.27	254.10	3638.35	↓ 6.88%	6345.44

* Results are listed in the original papers, – denotes none

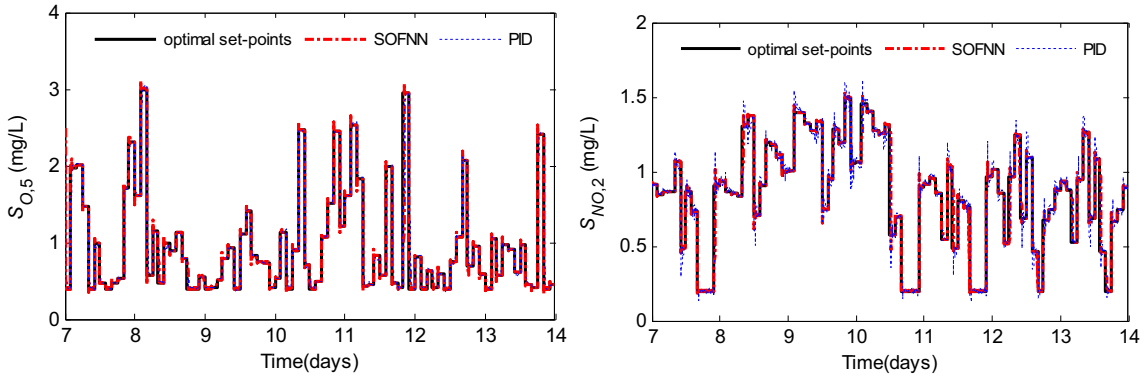

Fig. 16 Optimized set-points and online tracking control performance of S_O and S_{NO} in rain weather

Table 8 Comparison of EC and EQ with different optimal control strategies in rain weather

Weather	Method	AE	PE	EC		EQ
		(kWh/d)	(kWh/d)	(kWh/d)	Up/Down	(kg poll./d)
Rain	PID	3663.01	255.00	3918.01	–	7112.12
	SOOC [24]	3572.91*	266.74*	3839.65*	↓ 2.20%*	7215.23*
	DDAOC [16]	–	–	3741.69	↓ 4.50%*	–
	Hopfield [17]	–	–	3733.07	↓ 4.72%*	–
	RTO-NMPC [18]	3450.09*	268.70*	3718.19*	↓ 5.10%*	7596.60*
	NSGA-II [24]	3459.12*	265.34*	3724.46*	↓ 4.94%*	7368.96*
	MOPSO [27]	–	–	3687.63*	↓ 5.88%*	735336*
	MOEA/D-DE [38]	3389.03	269.01	3658.64	↓ 6.62%	7281.25
	MOEA/D-APS [34]	3411.77	278.99	3690.76	↓ 5.80%	7252.45
	MOEA/D-TPN [35]	3388.12	268.16	365628	↓ 6.68%	7310.70
	iMOEA/D [37]	3405.51	277.02	368253	↓ 6.01%	7260.33
	AMOEA/D	3301.44	344.26	3645.70	↓ 6.95%	7301.41

*Results are listed in the original papers, – denotes none

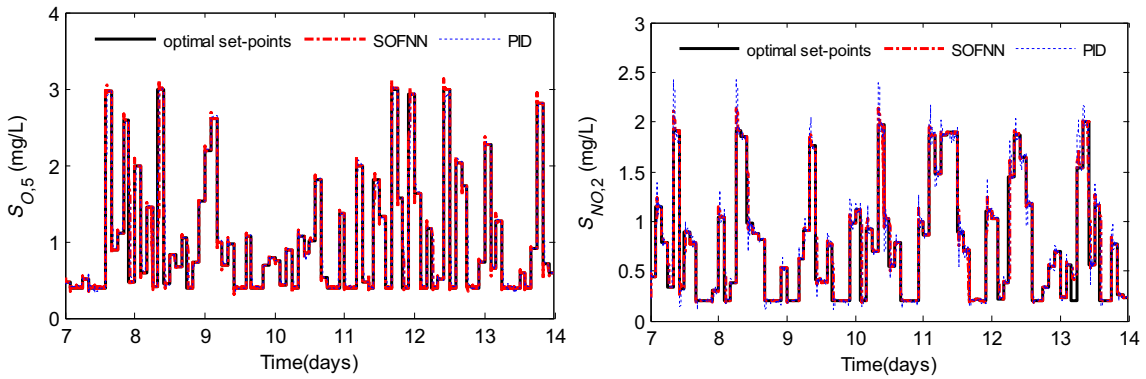

Fig. 17 Optimized set-points and online tracking control performance of S_O and S_{NO} in storm weather

Table 9 Comparison of EC and EQ with different optimal control strategies in storm weather

Weather	Method	AE	PE	EC		EQ
		(kWh/d)	(kWh/d)	(kWh/d)	Up/Down	(kg poll./d)
Storm	PID	3685.87	245.25	3931.12	–	6635.92
	SOOC [24]	3515.32*	337.18*	3852.50*	↓ 2.64%*	6815.03*
	DDAOC [16]	–	–	3717.26*	↓ 5.44%*	–
	Hopfield [17]	–	–	3683.06*	↓ 6.31%*	–
	RTO-NMPC [18]	3390.30	297.09*	3687.39*	↓ 6.20%*	7082.88*
	NSGA-II [24]	3425.43*	270.21*	3695.64*	↓ 5.99%*	7055.32*
	MOPSO [27]	–	–	366576*	↓ 6.75%*	7236.45*
	MOEA/D-DE [38]	3428.91	243.54	3672.45	↓ 6.58%	6805.57
	MOEA/D-APS [34]	3452.85	239.25	369210	↓ 6.08%	6766.92
	MOEA/D-TPN [35]	3441.48	223.89	3665.37	↓ 6.76%	6834.84
	iMOEA/D [37]	3449.20	218.14	3667.34	↓ 6.71%	6828.20
	AMOEA/D	3433.60	216.84	3650.44	↓ 7.44%	6853.80

*Results are listed in the original papers, – denotes none

PID controller, achieving a fast and high-precision tracking control under complex conditions.

In addition, the optimization effect of the AMOEA/D-based HMOOC strategy is compared with other optimal controllers: SOOC [25], DDAOC [16], Hopfield [17], RTO-NMPC [18], NSGA-II [25], MOPSO [27], MOEA/D-DE [38], MOEA/D-APS [34], MOEA/D-TPN [35], and iMOEA/D [37]. The results of AE, PE, EC, EQ and the average effluent parameters are given in Table 7, where influent represents the average influent parameters and PID denotes the constant control. From Table 7, in comparison with the PID constant control, the EC of AMOEA/D decreases by 6.88% in dry weather. Besides, the increase of EQ is smaller than that in SOOC and NSGAII. The experimental results indicate that AMOEA/D can not only ensure that the effluent parameters meet the standards, but also effectively reduce EC. In comparison with four MOEA/D variants, AMOEA/D presents an improved optimization effect, indicating that the proposed AMOEA/D with multiple adaptive strategies could obtain a set of high-quality optimal solutions. Compared with MOOC methods, the decrease of EC in SOOC is lower, and its energy-saving effect is less obvious. The increase of EQ in NSGAII is larger mainly due to the high effluent S_{NH}.

To further verify the adaptability of AMOEA/D in the complex working conditions, the experiments were carried out with the usage of data files of rain weather. The optimal control results of $S_{O,5}$ and $S_{NO,2}$ under the rain weather are presented in Fig. 16. It can be seen that the set-points of $S_{O,5}$ and $S_{NO,2}$ can be dynamically optimized by the AMOEA/D in the case of the inflow rate and component concentration with large disturbance. In addition, the SOFNN controller has high tracking accuracy and sound anti-interference ability.

Table 8 presents the comparative results of different optimal controllers in rain weather. Compared with the PID constant control, the EC of AMOEA/D decreases by 6.95% in rain weather. In addition, the EC of AMOEA/D is lower

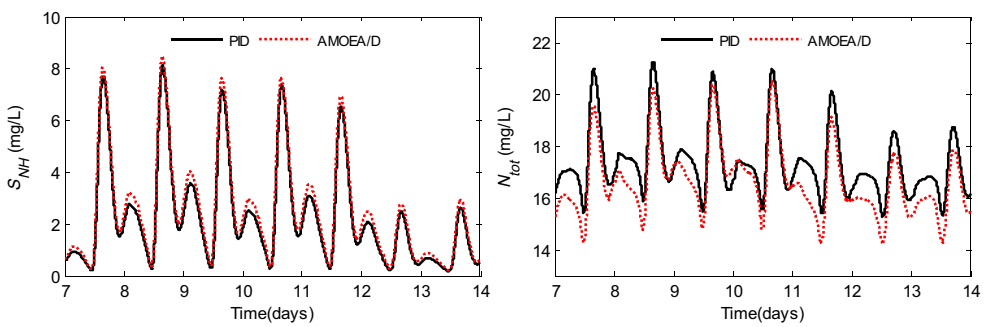

Fig. 18 Comparison of effluent parameters in dry weather

than that of NSGAII, MOPSO, and four MOEA/D variants, indicating that it can achieve a good energy-saving effect under complex conditions.

Figure 17 plots the optimized set-points of $S_{O,5}$ and $S_{NO,2}$ under the storm weather. The optimal results indicate that AMOEA/D can adjust the value of decision variables online regardless of the complex influent conditions. Table 9 presents the comparative results of different optimal controllers in storm weather. Compared with the PID constant control, the EC of AMOEA/D decreases by 7.44% in storm weather. In addition, the EC of AMOEA/D is lower than that of NSGAII, MOPSO, and four MOEA/D variants, indicating that it can achieve a good energy-saving effect under complex conditions.

6.4 Discussion

Figure 18 shows the variation of the effluent S_{NH} and N_{tot} in dry weather. In comparison with the PID constant control, the concentration of S_{NH} in the optimal control exhibits a slight increase in some optimal periods, and the concentration of N_{tot} shows a downward trend in general. In addition, BOD_5, COD and TSS remain unchanged, indicating that the HMOOC strategy can effectively reduce EC under the condition of ensuring the average effluent parameters to meet the standards. According to the analysis of the mechanism of the WWTP, S_{NH} and N_{tot} are a pair of parameters with competitive relationship, and the optimal balance between them can be achieved with the usage of the proposed AMOEA/D.

According to the further analysis of Table 7, S_{NH} slightly increases and N_{tot} decreases to a certain extent in AMOEA/D, whereas the other effluent parameters remain almost unchanged. These findings are consistent with the changes in the effluent parameters as shown in Fig. 18. It should be noted that all the comparison methods adopted in this study can ensure that the average effluent parameters satisfy the standard limits. According to the analysis of the removal rate of effluent contaminants, when compared with the influent parameters, the removal rates of the effluent S_{NH}, N_{tot}, BOD_5, COD and TSS can respectively reach 90.88%, 70.16%, 96.20%, 71.58%, and 93.66% by utilizing the AMOEA/D-based HMOOC strategy.

Furthermore, the average values of effluent S_{NH}, N_{tot}, BOD_5, COD and TSS in rain and storm weather are presented in Table 10. The results indicate that all the average value of effluent S_{NH}, N_{tot}, BOD_5, COD and TSS are remained within the limitations under complex weather conditions. The optimal control results shown in Figs. 15–18 and the performance indexes of different optimal controllers given in Tables 7–10 illustrate the effectiveness of the proposed AMOEA/D. The experimental results clearly indicate that the AMOEA/D-based HMOOC strategy can address the MOP in the WWTP under complex working conditions. In addition, the optimal results show that the objective functions constructed by the SOFNN approximator can meet the optimization needs of WWTP with high modeling accuracy. The multi-variable self-organizing controller can track the optimal set-points of $S_{O,5}$ and $S_{NO,2}$ with good stability and high control precision.

Table 10 Average effluent parameters of the proposed AMOEA/D in rain and storm weather

Limits/Weather	S_{NH}	N_{tot}	BOD_5	COD	TSS
	(mg/L)	(mg/L)	(mg/L)	(mg/L)	(mg/L)
Limits	4	18	10	100	30
Rain	3.65	16.14	2.88	47.18	13.34
Storm	3.88	16.30	2.85	47.69	13.55

7 Conclusion

In this paper, an HMOOC strategy based on the adaptive MOEA/D algorithm is proposed for the multiobjective optimization of EC and EQ in the WWTP. In fact, good optimal control performance mainly benefits from the following aspects. First, the HMOOC strategy, consisting of the SOFNN approximator, AMOEA/D algorithm and SOFNN controller, shows a good overall performance. The modeling accuracy of SOFNN approximator and the tracking precision of SOFNN controller could be enhanced by the self-organizing adjustment of the fuzzy rules and adaptive learning of network parameters, making them useful to the optimal control system. Second, for AMOEA/D, the AOS-based adaptive DE strategy could balance global exploration and local exploitation of the algorithm. The convergence and diversity of the Pareto solutions obtained by AMOEA/D could be improved. Third, based on the constraints handling mechanism, AMOEA/D can find more solutions with small violation values while performing the updating operation. Finally, by employing the two-phase optimization with auto-switching scheme, AMOEA/D can dig out more boundary solutions as well as effectively improve the quality of the candidate solutions, which is more suitable for the complicated MOPs in practical engineering. Future study is needed to fully explore the knowledge in the WWTP, and to construct an HMOOC strategy based on knowledge and processed data.

Acknowledgements The authors would like to thank the Editor-in-Chief, the Associate Editor and anonymous reviewers for their invaluable suggestions which have been incorporated to improve the quality of the paper. This work was supported in part by the National Science Foundation for Distinguished Young Scholars of China under Grant 61225016 and the State Key Program of National Natural Science of China under Grant 61533002.

References

1. Wan JF, Gu J, Zhao Q, Liu Y (2016) COD capture: a feasible option towards energy self-sufficient domestic wastewater treatment. Sci Rep 6(4):1–9
2. Oturan MA, Aaron JJ (2014) Advanced oxidation processes in water/wastewater treatment: principles and applications. A review. Crit Rev Environ Sci Technol 44(23):2577–2641
3. Santín I, Pedret C, Vilanova R, Meneses M (2015) Removing violations of the effluent pollution in a wastewater treatment process. Chem Eng J 279(11):207–219
4. Judd SJ (2016) The status of industrial and municipal effluent treatment with membrane bioreactor technology. Chem Eng J 305(12):37–45
5. Åmand L, Carlsson B (2012) Optimal aeration control in a nitrifying activated sludge process. Water Res 46(7):2101–2110
6. Wahab NA, Katebi R, Balderud J (2009) Multivariable PID control design for activated sludge process with nitrification and denitrification. Biochem Eng J 45(3):239–248
7. Song X, Zhao Y, Song Z (2012) Dissolved oxygen control in wastewater treatment based on robust PID controller. Int J Modell Identif Control 15(4):297–303
8. Holenda B, Domokos E, Redey A, Fazakas J (2008) Dissolved oxygen control of the activated sludge wastewater treatment process using model predictive control. Comput Chem Eng 32(6):1270–1278
9. Belchior CAC, Araújo RAM, Landeck JAC (2012) Dissolved oxygen control of the activated sludge wastewater treatment process using stable adaptive fuzzy control. Comput Chem Eng 37:152–162
10. Qiao JF, Zhang W, Han HG (2016) Self-organizing fuzzy control for dissolved oxygen concentration using fuzzy neural network. J Intell Fuzzy Syst 30(6):3411–3422
11. Hreiz R, Latifi MA, Roche N (2015) Optimal design and operation of activated sludge processes: state-of-the-art. Chem Eng J 281(12):900–920
12. Ostace GS, Baeza JA, Guerrero J, Guisasola A (2013) Development and economic assessment of different WWTP control strategies for optimal simultaneous removal of carbon, nitrogen and phosphorus. Comput Chem Eng 53(6):164–177
13. Santin I, Pedret C, Vilanova R (2015) Applying variable dissolved oxygen set point in a two level hierarchical control structure to a wastewater treatment process. J Process Control 28(4):40–55
14. Guerrero J, Guisasola A, Vilanova R, Baeza JA (2011) Improving the performance of a WWTP control system by model-based setpoint optimisation. Environ Modell Softw 26(4):492–497
15. Machado VC, Gabriel D, Lafuente J, Baeza JA (2009) Cost and effluent quality controllers design based on the relative gain array for a nutrient removal WWTP. Water Res 43(20):5129–5141
16. Qiao JF, Bo YC, Chai W, Han HG (2013) Adaptive optimal control for a wastewater treatment plant based on a data-driven method. Water Sci Technol 67(10):2314–2320
17. Han G, Qiao JF, Han HG, Chai W (2014) Optimal control for wastewater treatment process based on Hopfield neural network. Control Decis 29(11):2085–2088
18. Vega P, Revollar S, Francisco M, Martín JM (2014) Integration of set point optimization techniques into nonlinear MPC for improving the operation of WWTPs. Comput Chem Eng 68:78–95
19. Dai HL, Chen WL, Lu XW (2016) The application of multi-objective optimization method for activated sludge process: a review. Water Sci Technol 73(2):223–235
20. Hakanen J, Sahlstedt K, Miettinen K (2013) Wastewater treatment plant design and operation under multiple conflicting objective functions. Environ Model Softw 46(4):240–249
21. Sweetapple C, Fu G, Butler D (2014) Multi-objective optimisation of wastewater treatment plant control to reduce greenhouse gas emissions. Water Res 55(2):52–62
22. Hreiz R, Roche N, Benyahia B, Latifi MA (2015) Multi-objective optimal control of small-size wastewater treatment plants. Chem Eng Res Des 102(7):345–353
23. Chen WL, Lu XW, Yao CH (2015) Optimal strategies evaluated by multi-objective optimization method for improving the performance of a novel cycle operating activated sludge process. Chem Eng J 260(9):492–502
24. Zhang R, Xie WM, Yu HQ, Li WW (2014) Optimizing municipal wastewater treatment plants using an improved multi-objective optimization method. Bioresour Technol 157(2):161–165
25. Qiao JF, Zhang W (2016) Dynamic multi-objective optimization control for wastewater treatment process. Neural Comput Applic 28(10):1–11
26. Qiao JF, Hou Y, Zhang L, Han HG (2018) Adaptive fuzzy neural network control of wastewater treatment process with multiobjective operation. Neurocomputing 275:383–393
27. Han HG, Zhang L, Liu HX, Qiao JF (2018) Multiobjective design of fuzzy neural network controller for wastewater treatment process. Appl Soft Comput 67:467–478
28. Zhang QF, Li H (2007) MOEA/D: a multiobjective evolutionary algorithm based on decomposition. IEEE Trans Evol Comput 11(6):712–731
29. Li K, Zhang QF, Kwong S, Li MQ, Wang R (2014) Stable matching-based selection in evolutionary multiobjective optimization. IEEE Trans Evol Comput 18(6):909–923
30. Wu MY, Li K, Kwong S, Zhou Y, Zhang QF (2017) Matching-based selection with incomplete lists for decomposition multi-objective optimization. IEEE Trans Evol Comput 21(4):554–568
31. Zhao SZ, Suganthan PN, Zhang QF (2012) Decomposition-based multiobjective evolutionary algorithm with an ensemble of neighborhood sizes. IEEE Trans Evol Comput 16(3):442–446
32. Wang ZK, Zhang QF, Zhou AM, Gong MG, Jiao LC (2016) Adaptive replacement strategies for MOEA/D. IEEE Trans Cybern 46(2):474–486
33. Qi YT, Ma XL, Liu F, Jiao LC, Sun JY, Wu JS (2014) MOEA/D with adaptive weight adjustment. Evol Comput 22(2):231–264
34. Yang SX, Jiang SY, Jiang Y (2017) Improving the multiobjective evolutionary algorithm based on decomposition with new penalty schemes. Soft Comput 21(16):4677–4691
35. Jiang SY, Yang SX (2016) An improved multiobjective optimization evolutionary algorithm based on decomposition for complex Pareto fronts. IEEE Trans Cybern 46(2):421–437
36. Wang ZK, Zhang QF, Li H, Ishibuchie H, Jiao LC (2017) On the use of two reference points in decomposition based multiobjective evolutionary algorithms. Swarm Evol Comput 34:89–102
37. Ho-Huu V, Hartjes S, Visser HG, Curran R (2018) An improved MOEA/D algorithm for bi-objective optimization problems with complex Pareto fronts and its application to structural optimization. Expert Syst Appl 92:430–446
38. Li H, Zhang QF (2009) Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II. IEEE Trans Evol Comput 13(2):284–302
39. Li K, Fialho A, Kwong S, Zhang QF (2014) Adaptive operator selection with bandits for a multiobjective evolutionary algorithm based on decomposition. IEEE Trans Evol Comput 18(1):114–130
40. Sallam KM, Elsayed SM, Sarker RA (2017) Landscape-based adaptive operator selection mechanism for differential evolution. Inform Sci 418:383–404
41. Lin QZ, Liu ZW, Yan Q, Du ZH, Coello CAC, Liang ZP, Wang WJ, Chen JY (2016) Adaptive composite operator selection and parameter control for multiobjective evolutionary algorithm. Inform Sci 339:332–352

42. Lin QZ, Ma YP, Chen JY, Zhu QL, Coello CAC, Wong KC, Chen F (2018) An adaptive immune-inspired multi-objective algorithm with multiple differential evolution strategies. Inform Sci 430:46–64
43. Qiao JF, Zhou HB (2017) Prediction of effluent total phosphorus based on self-organizing fuzzy neural network. Control Theory Applic 34(2):224–232
44. Qiao JF, Zhou HB (2018) Modeling of energy consumption and effluent quality using density peaks-based adaptive fuzzy neural network. IEEE/CAA J Automatica Sinica 5(5):968–976
45. Zhou HB (2017) Dissolved oxygen control of the wastewater treatment process using self-organizing fuzzy neural network. CIESC J 68(4):1516–1524
46. Jeppsson U, Pons MN (2004) The COST benchmark simulation model-current state and future perspective. Control Eng Pract 12(3):299–304
47. Santín I, Pedret C, Vilanova R, Meneses M (2016) Advanced decision control system for effluent violations removal in wastewater treatment plants. Control Eng Pract 49:60–75
48. Jan MA, Khanum RA (2013) A study of two penalty-parameterless constraint handling techniques in the framework of MOEA/D. Appl Soft Comput 13(1):128–148
49. Zhu QL, Lin QZ, Chen WN, Wong KC, Coello CAC, Li JQ, Zhang J (2017) An external archive-guided multiobjective particle swarm optimization algorithm. IEEE Trans Cybern 47(9):2794–2808

Hongbiao Zhou received the B.E. degree in electrical engineering and automation from the Huaiyin Institute of Technology, Huai'an, China, in 2003, and the M.E. degree in control theory and control engineering from Lanzhou University of Technology, Lanzhou, China, in 2009. He is currently working toward the Ph.D. degree with the Beijing University of Technology, Beijing, China. His current research interests include neural networks, multi-objective optimization, and process control systems.

Junfei Qiao received the B.E. degree and M.E. degree in control engineering from Liaoning Technical University, Fu'xin, China, in 1992 and 1995, respectively, and the Ph.D. degree from Northeast University, Shenyang, China, in 1998. From 1998 to 2000, he was a Postdoctoral Fellow with the School of Automatics, Tianjin University, Tianjin, China. He then joined the Beijing University of Technology, Beijing, China, where he is currently a Professor and the Director of the Intelligence Systems Laboratory. His research interests include neural networks, intelligent systems, self-adaptive/learning systems, and process control systems.