Evolutionary Optimization with Adaptive Surrogates and its Application in Crude Oil Distillation

Xuhua Shi, Chudong Tong
Faculty of Electrical Engineering & Computer Science
Ningbo University
Ningbo, China
shixuhua@nbu.edu.cn

Li Wang
School of electrical and Electronic Engineering
Shanghai institute of technology
Shanghai, China
wangli@sit.edu.cn

***Abstract*—Surrogate modelling and model management are key points for evolutionary optimization of chemical processes. This paper proposes an evolutionary algorithm with the help of adaptive surrogate functions (EASF), in which approximate models' establishment and management are combined to search the optimal result. To construct an appropriate surrogate model, a new hybrid modelling framework with adaptive Radial Basis Functions (RBF) (ARBF) is put forward. Different from most neural network modelling methods, ARBF is able to adaptively adjust the sample size by current approximation errors to effectively take into account the tradeoff between approximation accuracy and sample size. For model management, an approximation error fuzzy control strategy (AEFCS) is introduced. AEFCS in combination with ARBF can effectively perform exploratory and exploitative search in the evolutionary optimization. The superiority of EASF is demonstrated by the simulation results on three benchmark problems. To illustrate the performance of EASF further, it is employed to optimize the operating conditions of crude oil distillation process, and satisfactory results are obtained.**

Keywords—evolutionary optimization; surrogate modelling; model management; process optimization

I. INTRODUCTION

In the applications of crude oil distillation unit (CDU) modelling and optimization, the behaviors of a CDU are usually highly nonlinear and strongly coupled, and the mathematical expressions required to describe them are often complex [1]. It is time-consuming to figure out the optimal working conditions. Specific simulation software tools like Aspen Plus and Hysys have been developed to model and simulate the processes, which however, are black boxes to users and not convenient to access to the integrated system. To balance of the computational costs and the fidelity of the simulation, a computationally cheaper approximation expression, often called surrogate model can be used to fit the underlying system.

However, constructing an accurate surrogate model while obtaining optimal results with reduced computation cost remains a challenge. Typical promising surrogate models are Polynomial Regression, Gaussian Process, Multivariate Regression Splines, and Radial Basis Functions (RBF). Jin *et al.* [2] compared these models using multiple performance criteria, and showed that RBF is the best one in terms of its average accuracy and robustness. Furthermore, RBF can fit any complex nonlinear functions given sufficient complexity of the trained network. Recently, much work has been published on its applications for chemical process optimization. Yan *et al.* employed RBF to model the aromatic hydrocarbon isomerization (AHI) process based upon practical observation data and use chaos-genetic algorithm to optimize the operational condition of AHI [3]. Assisted by the RBF distillation surrogate model, Lu *et al*. proposed an on-line multivariable predictive control strategy to obtain more profit [4]. Liau *et al.* [5] , Motlaghi *et al.* [6] developed RBF models regressed from plant measurements, distillation simulations to build the distillation models, and using evolution algorithm to optimize the product yields. Ismail and Selen[7] developed RBF thermo dynamic models to reduce the computational cost, although large amount of data was needed. Similar applications can be found in [8-10].

In addition, RBF models offer other advantages like powerful function approximation and clear mathematical expression, etc. Like other approximate models, RBF models have some disadvantages: 1) the quality of the model depends on the quality of the data used during modelling. 2) the accuracy of the model is somewhat related to the size of samples, usually the bigger the sample size is, the higher the accuracy of the model will be.

The challenges for evolutionary optimization of CDU include : 1) how to obtain the qualified samples, 2) how to achieve an appropriate tradeoff between the sample size and approximate accuracy, 3) how to use the approximate models. So, in section II, we will provide a brief introduction to CDU processes, followed by Section III, discussing some key points for approximation modelling and evolutionary optimization and proposing an Adaptive Surrogate Function Evolution approach. Section IV presents the results on the benchmark problems as well as a CDU application. The conclusions and future research directions are given in Section V.

II. OPTIMIZATION OF CDU

A. *Description of CDU*

CDU is one of the important units of the refinery atmospheric and vacuum distillation process. Consider a CDU

consisting of a fifty stage main distillation column, one condenser, four side strippers and four pump-arounds, as illustrated in Fig. 1. The preheated crude oil enters the column at the 35th stage. The distillation system processes about 0.25 m3/s of light crude oil into several products, namely Naphtha, Special Cut Naphtha(SCN), Superior Kerosene(SK), light gas oil(LGO), heavy gas oil (HGO) and residue etc. The distillation employs steam at 250□as a stripping agent. Steam flow rates for the main column is about 800 kmol/h, Reflux ratio in the main column equals 4.17 and operating pressure in the distillation unit is 2.5 bar. The CDU tower has four middle circulation refluxes, for example, the Top Circulation Reflux (Top CR) is withdrawn from the 2th stage through a heat exchanger and is cooled to 70□, then return to the 1th tower stage.

There are two reasons for setting the circulating refluxes in the distillation tower, one is to get uniform vapor - liquid distribution, so as to improve the separation ability of the tower. The second is to recover the tower waste heat, to increase the heat utilization efficiency of the tower. The circulation reflux of the distillation column is incorporated with the heat exchange network, and heat is fully exchanged with the crude oil, so that the crude oil is fully preheated before entering the furnace to reduce the fuel consumption and the production cost.

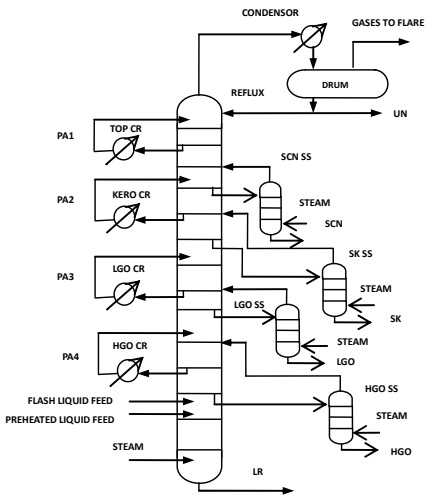


Fig. 1 Crude Oil Distillation Column

B. Description of the Optimization Problem

Different optimization problems involve different objectives. The optimization objective chosen here is to maximize the total distillation heat supply. Additionally, the constraints are mainly considered to be specific products performance indicators. Below is the expression of the involved problem.

$$\begin{aligned} &max\ J = \sum_{i=1}^{4} E_{n_i}(\mathbf{u}) \\ &s.t. \begin{cases} \alpha \le x_i^{dist}(\mathbf{u}) \le \beta \\ E_{ni} = Distillation\ MESH(\mathbf{u}) \end{cases}, i = 1,2,3,4 \end{aligned} \tag{1}$$

Where $E_{n_i}, i = 1,2,3,4$ represent the distillation heat supply of Top Circulating Reflux (TCR), Kerosene Circulating Reflux (KCR), LGO Circulating Reflux(LCR), and HGO Circulating Reflux (HCR), which are related with the flow rate of TCR ,KCR, LCR, and HCR, $\mathbf{u}$ is the vector of the decision variables which are selected based on the knowledge of distillation heat supply. TABLE I shows the base case values of the selected decision variables, which include products flow rate of process streams, reflux ratios of pump-arounds. The constraints are distillation product property expressed in terms of T5% and T95% on the TBP (true boiling point) curve and distillation's four balance equations (*MESH* equations). They are illustrated in Eq.(1).

TABLE I. BASE CASE VALUES OF THE DECISION VARIABLES

Item	Lower bound	Upper bound	Base case
UN flow rate(m3/s)	0.0150	0.0210	0.0180
SCN flow rate(m3/s)	0.0200	0.0227	0.0204
SK flow rate(m3/s)	0.0193	0.0223	0.0213
LGO flow rate(m3/s)	0.0390	0.0446	0.0406
HGO flow rate(m3/s)	0.0106	0.0160	0.0124
PA1 reflux ratio	0.1	0.9	0.25
PA2 reflux ratio	0.1	0.9	0.25
PA3 reflux ratio	0.1	0.9	0.25
PA4 reflux ratio	0.1	0.9	0.25

III. EVOLUTIONARY OPTIMIZATION WITH ADAPTIVE SURROGATE FUNCTIONS

A. Challenges

Since the early introduction of surrogate models into chemical processes [11], much progress has been made in the development of the optimization [12]. However, the use of surrogate models has brought new challenges due to more complex structure and relation between the decision variables [13].

- Data sources: Selection of data samples is mainly determined by the optimization types employed. Generally, the types of industrial process optimization mainly include real time optimization and operating direction optimization. For example, if the samples are applied for the former, it would be reasonable to collect the data from plant real time measurements, laboratory on-line analysis, or soft sensor results [14-15]. In contrast, if they are applied for the latter, it is convenient to collect the data from the rigorous simulations [1,16]. No matter what kinds of samples are used, data pre-procession and pre-calibration must be taken into account to reduce the impact of measurement errors [17].
- Sampling: A successful implementation of sampling is to obtain well-distributed and complete sampling points[18]. Generally, sampling methods can be categorized into one-shot and sequential design [19]. The latter is superior to the former is that new sampling points in the next iteration are generated using the information of the approximated models in the current iteration. Since highly nonlinear characteristics of the

complex chemical process, sequential design sampling is very popular[2,8,9]. In terms of the accuracy of the surrogate model, it depends on the number of samples to a certain extent. So there exists a tradeoff between the sample sizes and approximation accuracy.

- How to use the surrogate models: Actually, the surrogate model can be applied in almost all operations of evolutionary period, such as population initialization, offspring reproduction, pre-selection, and fitness evaluations [20-22]. Now days, various surrogate model management approaches have been emerged yet, such as, global and local surrogate assisted framework [23], two-layer surrogates management [24], and multiple operator search strategy [25], among many others. However, a challenge might be still encountered on how to trade off the optimization accuracy with computational costs, that is, how to obtain a satisfactory result while using a minimum number of consuming model evaluations.

B. The proposed method

Confronted with highly complicated chemical process, one important thing is to choose an appropriate sampling method. For one-shot design, the samples used for building an approximate model are generated before the execution of the experiment. So, inflexibility in learning the complicated behavior of the underlying process is usually suffered. If the obtained approximate model isn't able to meet the performance requirements, then a new sampling experiment will be re-conducted.

To overcome drawbacks mentioned above, an evolutionary algorithm assisted by adaptive surrogate functions (called EASF) is put forward. Fig.2 illustrates the framework of the proposed method. The main components of the proposed framework include an Optimizer, an Adaptive Sampling System and a Model Manager. During the evolution, the original models and approximate models' management is performed by the Manager, while the Optimizer corporation with the Manager conducts the evolutionary optimization task. The Adaptive Sampling system synchronized with the evolutionary optimization implements data points sampling, rigorous model simulation, surrogate model training and approximation accuracy evaluation. Firstly, a small number of training samples with uniform distribution in the search space are generated to construct an initial RBF model, which is expressed by the following equation:

$$f(X)=\sum_{i=1}^{N}\omega_i \exp\left\{\frac{-\|X-M_i\|^2}{2\theta_i^2}\right\} \qquad (2)$$

where N is the number of training samples; ω_i, M_i and θ_i are coefficients of i th Gaussian function. At each iteration, the population at Optimizer is generated or updated using traditional evolution operations, such as reproduction, pre-selection, and fitness evaluations. The individuals' evaluation whether by rigorous models or surrogate models are controlled by the Model Manager. At each step of the evolution, the current approximate models are evaluated by the Model Evaluator, if the approximate accuracy isn't able to meet the performance requirement, new data points will be generated by the Rigorous Simulator, then, they will be added to the training data set and the RBF model will be retrained and renewed. Notice that, in EASF, new sampling points for next iteration will be generated using the information of the constructed models and design points in the current iteration, and that, EASF performs data sampling and optimization simultaneously, can effectively tradeoff the sample size and optimization accuracy. The pseudo code of the main components of the proposed EASF is listed in Algorithm 1.

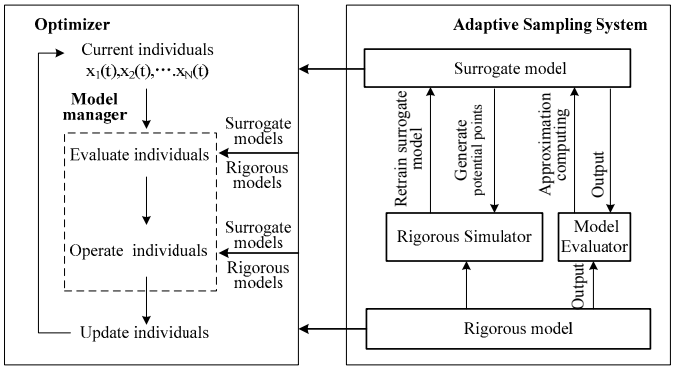


Fig.2 An illustration of EASF

Algorithm 1. The main framework of the proposed EASF

1: **Input:** the maximal number of generations t_{max}, size of population N_p, initial Latin-hypercube samples N_0, allowed approximate error δ.
2: **Output:** final optimal result
3: /* Initialization */
4: Initialization Population: Randomly generate N_0 size of the initial population $\mathbf{P}_0=\{\mathbf{p}_1,\mathbf{p}_2,\cdots,\mathbf{p}_{N0}\}$;
5:Generate the initial training set by Latin-hypercube sampling and train an initial RBF network;
6: Set t=0;
7: /*Main Loop*/
8: **while** $t < t_{\max}$ **do**
9: Do model management operation: Evaluate individuals with surrogate model S, rigorous model R.
10: Q_t =offspring reproduction (P_t).
11: $P_t = Q_t \cup P_t$
12: P_{t+1} =Selection (t,P_t)
13: **while** approximate accuracy > δ **do**
14: Do Adaptive sampling operation
15: **end while**
16: **end while**

C. Adaptive Sampling Strategy

The choice of data samples is of great importance to the success of the surrogate modelling. As illustrated in Fig. 3, the proposed strategy implements data sampling, data renewing and approximate models training in an iterative way. Following are some introductions.

At each iteration, the Model Evaluator will assess the accuracy of the approximate models, if the approximation error is less than the allowed one, i.e., $\varepsilon_j < \delta_j, j = 1,2,\cdots ns$,

then the constructed model is satisfied, and the approximate model is final obtained. Here, $\varepsilon = \sqrt{\sum_{i=k-\lambda}^{k}(S(\mathbf{x}_i) - R(\mathbf{x}_i))^2}$, δ refers to the allowed maximum approximate error, $S(\mathbf{x}), R(\mathbf{x})$ denote to the outputs of surrogate model and rigorous model respectively, $\mathbf{x}$ is the decision variables. k is the current cycle of the generations, $\lambda < k$ refers to evaluation length for the approximate model, ns is the surrogate model numbers. Conversely, if the approximate model isn't able to meet the desired accuracy, then the model should be renewed. It was demonstrated in [26] that putting the potential data points of the surrogate model to the sample set and renewing the model can accelerate the convergence speed and meanwhile enhance the local searching ability. In the light of this, the strategy of new data points adding to the sample set in our proposed method is expressed as follows:

$$\begin{cases} Sp = Sort_{\max}^{m_j}(S_j(\mathbf{x})) \bigcup Sort_{\min}^{n_j}(S_j(\mathbf{x})), j = 1,2,\cdots ns \\ s.t. \quad \mathbf{x} \in R^n \\ \quad\quad d(\mathbf{x}) \geq \sigma \end{cases} \tag{3}$$

where, $S_j, j = 1,2,\cdots,ns$ are the surrogate models being constructed. $Sort_{\max}^{m_j}$ and $Sort_{\min}^{n_j}$ represent m_j maximum and n_j minimum potential points, d is the distance between two sample points, σ is the threshold of the distance.

$$\sigma = \min(((x_i - x_j)^T (x_i - x_j))^{1/2}), \quad i, j = 1,2,\cdots,N, i \neq j \tag{4}$$

where x_i, x_j represent the sample point in the data set, N is the number of samples, which will be increased during the iteration. It should be noticed that the determination of the new potential sample size is adaptively controlled by the approximation error in expression $m_j, n_j = \left\lceil \frac{N_0 \cdot (\varepsilon_j - \delta_j)}{\varepsilon_0 - \delta_j} \right\rceil$, where $N_0, \delta, \varepsilon_0$ *and* ε represent the number of initial Latin-hypercube samples, the allowed approximate error, initial approximation error and current approximation error respectively.

From eq. (3), we can see that in the early stages of modelling, when there are fewer samples, the surrogate models may be less accurate, not enough to reflect the information of the whole modelling space. Hope that through the exploration of new sampling points, the global approximation accuracy of the updated model can be improved. Therefore, the distance threshold σ should be appropriate to take a little bigger to exploration the domain space. While in the late stages, with the new sampling points continuously inputting to the modelling data set, and continuously updating the surrogate model, therefore the global approximation of the model can be improved. It is appropriate to reduce the distance threshold σ to exploitation the domain. The distance threshold given in eq. (4) can be adaptively adjusted with the increasing number of the samples. It is an adaptive mixture of explorative - exploitative searching method. Following is the algorithm framework.

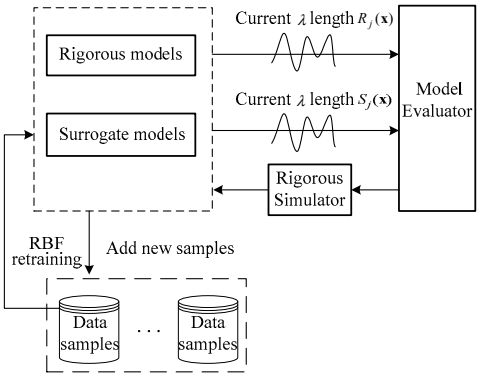


Fig.3 Illustration of adaptive sampling

Algorithm 2. The main framework of Adaptive sampling operation

1: **Input:** Initial Latin-hypercube samples $\boldsymbol{N}_0$, allowed approximate error δ, maximum sizes of new points n_m, surrogate model numbers ns.
2: **Output:** Surrogate model S;
3: **for** surrogate model $S_j, j = 1,2\cdots, ns$
4: Compute approximate error ε_j;
5: **while** approximate accuracy > δ_j **do**
6: Compute new sample numbers m_j, n_j;
7: Search m_j maximum and n_j minimum potential points in S_j model;
8: Do rigorous model simulation and generate new data samples;
9: Retrain the surrogate model;
10: **end while**
11:**end for**

D. Model Management Strategies

Model management is about how to manipulate of the approximate models to ensure the search process converges to a reasonable solution of the original problem. It is a great challenge when the approximate function has false optima. In [25], a detailed analysis was given on individual-based and generation-based approximate model management. Further, a framework of individual-based management was proposed by combining the original fitness function with the approximate fitness function.

While, for chemical process optimization, with the constant requirement of the high fidelity surrogate models, usually more than one surrogate models need to be constructed. To manage multiple surrogate models and ensure the correct convergence of the evolution, in this paper, an individual-based approximation error fuzzy control strategy (AEFCS) using heuristic fuzzy rules is proposed. Following is the expression for the rules :

for generation t:**if** ($\psi(\xi_1), \psi(\xi_2), \cdots \psi(\xi_{ns})$) **then** R ; **end for**

where;

$$\xi_i = 1 - (1 + (\varepsilon_i / \delta_i)) \cdot e^{-(\varepsilon_i / \delta_i)}, i = 1,2,\cdots,ns \tag{5}$$

$\xi_i, i = 1,2,\cdots,ns$ are input variables, which represent the error information of surrogate models. $\varepsilon_i, \delta_i, i = 1,2,\cdots,ns$ refer to the current approximation error and allowed maximum approximation error respectively. $\psi(x)$ is fuzzy membership operator, such as Gaussian, Sigmoid, and Trapezoidal operators, etc. $R = \gamma_1 \cup \gamma_2 \cdots \cup \gamma_{ns}$ are outputs, which represent the model management results. If $\gamma_1 = 1$ denotes that the corresponding individual should be evaluated using original fitness, vice versa.

IV. Results on Benchmark Problems and Real-world Applications

A. Benchmark problems

To verify the feasibility of the proposed approach, three test functions as benchmark functions have been considered in this work: 2D Griewank function on domain $[-10,10]^2$, 2D Rastrigin function on domain $[-5,5]^2$, and 2D Ackley on domain $[-5,5]^2$. The functions of Griewank with 5D and 10D on domain $[-10, 10]^5$ and $[-10, 10]^{10}$ are also used to evaluate the performance in high dimensions. For sampling operation, considering more complicated for 5D and 10D Griewank functions, the initial training sample size is set to 40 to train a three layer RBF neural networks. For the other 2D functions, the initial sample size is set to 20.

In this work, we adopt the Differential Evolution Algorithm (DEA) as the optimizer. In the simulations, the number of particles is set to 100, the scaling factor is set to 0.5, and the crossover probability is 0.9, model evaluation length λ is 6. The maximal number of generations t is set to 200. In model management, Triangular and Trapezoidal membership functions are employed for $\psi(\xi_1) \in \{S, M, L\}$ with following fuzzy rules:

if $\psi(\xi_1) = S$,then $\gamma_1 = 0$;

if $\psi(\xi_1) = M$,then $\begin{cases} for \ \ t < 20, \gamma_1 = 0 \\ \text{else}, \gamma_1 = 1 \end{cases}$;

if $\psi(\xi_1) = L$,then $\gamma_1 = 1$;

Simulations are carried out on the above functions for three different cases: 1) evolution with the original function only (ORG-only), 2) evolution with surrogate function only(SUR-only), and 3) evolution with EASF. For comparing the results, the following performance criterions are adopted:

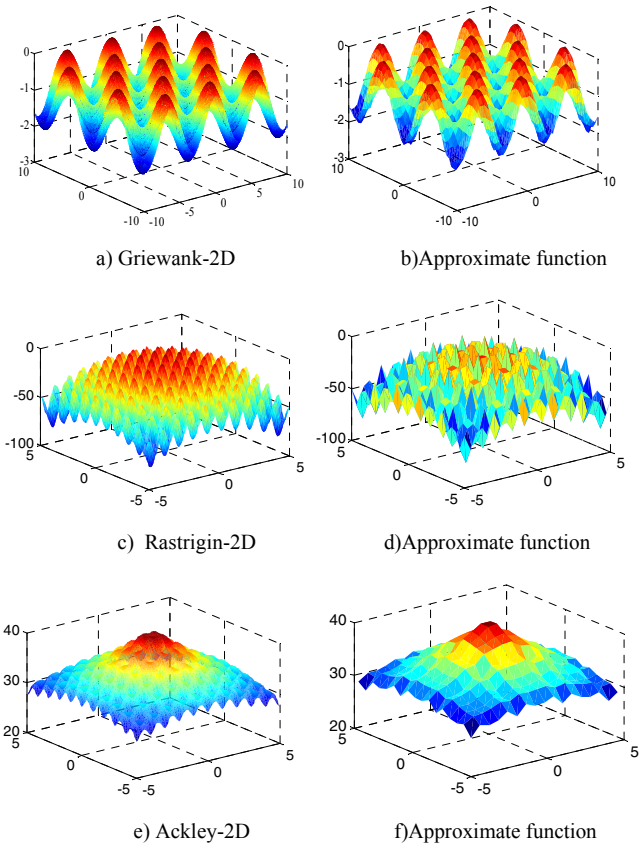
$$J_1 = \sum_{i}^{n_z} \left| f_i(\mathbf{x}_i) - \hat{f}_i(\mathbf{x}_i) \right| \tag{6}$$

$$J_2 = MAX_{i=1}^{n_z} \left| f_i(\mathbf{x}_i) - \hat{f}_i(\mathbf{x}_i) \right| \tag{7}$$

$$\phi = \frac{1}{T} \sum_{i=1}^{T} \theta_i \tag{8}$$

$$\Delta\varepsilon = \frac{1}{T} \sum_{i=1}^{T} \left| y_{min} - \hat{y}_{min} \right| \tag{9}$$

where $f(\mathbf{x})$, $\hat{f}(\mathbf{x})$ are the output of the assumptive process, the output of the RBF respectively. n_z is the iteration number of the final surrogate model. ϕ is T runs average standard deviation of DEA evolution result. y_{min} and $\hat{y}_{min}$ are T runs average minimum and computational minimum of the original function respectively. Graphical representations of three 2D challenge functions and their responses of the final surrogate model are shown in Fig. 4. TABLE II lists the optimization results of ORG-only, SUR-only and EASF.



a) Griewank-2D b)Approximate function

c) Rastrigin-2D d)Approximate function

e) Ackley-2D f)Approximate function

Fig.4 Approximate functions

From the simulation results, we can see that the proposed sampling method can mainly capture the behavior of the multi-modal functions, as long as more initial sample points are provided for high dimensional functions. TABLE II gives some optimization results for three approaches. In the 30 runs, for 2D multi-modal functions, when using proposed adaptive sampling, the cumulative error J_1 ranges from 1.23e+2 to 1.54e+2, the maximum error J_2 is from 1.97 to 2.85, and the average sampling points are about 1163 with the control of the allowed maximum approximate error. The results show that the proposed sampling method can effectively capture the behavior of the functions as illustrated in Fig. 4. The results are slightly inferior on high dimensional functions. As for evolution, 20000 calls of the original function are made on average in 200 generations for a population size of 100 with the maximum deviation from 0.45 to the minimum deviation of 0.08. The frequency at which the original function is called can be greatly reduced by using model management strategy. When the individual-based approximation error fuzzy control strategy (AEFCS)

is used, the original function called frequency are reduced from average 9544 to 8768 for high dimensional functions, and from average 6524 to 4464 for 2D functions with similar evolution accuracy compared with the results of evolution with original function only method. While using SUR-only, the average deviation is much bigger.

TABLE II . OPTIMIZATION RESULTS FOR THREE APPROACHES

Methods	Functions	δ	J_1	J_2	n_z	Initial points	Sampling points	ϕ	$\Delta\varepsilon$	Number of calls(AVG.)
ORG-only	Griewank-10D		—	—		—	—	2.15e-2	0.45	20000
	Griewank-5D		—	—		—	—	2.29e-2	0.28	20000
	Griewank-2D		—	—		—	—	2.31e-2	0.02	20000
	Rastrigin-2D		—	—		—	—	2.71e-2	0.08	20000
	Ackley-2D		—	—		—	—	2.44e-2	0.05	20000
SUR-only	Griewank-10D	0.1	2.73e+2	3.23	112	40	1621	8.37e-2	3.35	—
	Griewank-5D	0.1	2.49e+2	3.11	101	40	1587	5.78e-2	2.98	—
	Griewank-2D	0.01	1.23e+2	2.16	86	20	1210	3.91e-2	0.92	—
	Rastrigin-2D	0.02	1.37e+2	2.85	91	20	1109	1.31e-2	7.18	—
	Ackley-2D	0.02	1.31e+2	2.31	88	20	1408	6.44e-2	9.08	—
EASF	Griewank-10D	0.2	2.31e+2	3.91	107	40	1568	3.11e-2	0.57	9544
	Griewank-5D	0.2	2.79e+2	3.66	105	40	1467	3.32e-2	0.48	8768
	Griewank-2D	0.01	1.33e+2	1.97	88	20	1026	2.75e-2	0.02	4928
	Rastrigin-2D	0.05	1.47e+2	2.21	82	20	1012	1.01e-2	0.12	6524
	Ackley-2D	0.05	1.54e+2	2.72	78	20	1218	2.47e-3	0.16	4464

B. Application to CDU Optimization

CDU sample points are generated by simulation of its rigorous model, a total of 20 rigorous simulations in Aspen PLUS are performed to obtain the initial samples, which are generated using Latin hypercube sampling technique. For each converged simulation in ASPEN PLUS, MATLAB saves the necessary information to build the initial RBF model. Once the samples are obtained, the proposed adaptive modelling strategy is applied to obtain the reduced model. For the complexity of the CDU system, here, we construct four reduced parallel models corresponding to the optimization problem presented in (1). These models are reflux heat load model, Naphtha-Kerosene separation accuracy model, Kerosene-LGO separation accuracy model and LGO-HGO separation accuracy model as shown in TABLE III. For this optimization problem, the decision variables are listed in TABLE I, and the maximal number of generations is set to 60 also using DEA. To verify the feasibility of the proposed approach for CDU application, three different model management approaches as we did in the numerical experiments are applied in the CDU optimization. For CDU multiple model management, following are fuzzy rules:

if $\psi(\xi_i) = S$,then $\gamma_i = 0$, $i = 1,2,3,4$;

if $\psi(\xi_i) = M$,then

$$\begin{cases} for \;\; t < 10 \;\; and \;\; \psi(\xi_j) \; not \; 'L', then \;\; \gamma_i = 0 \\ else, \gamma_i = 1 \end{cases}, i \neq j;$$

if $\psi(\xi_i) = L$,then $\gamma_i = 1$, $i = 1,2,3,4$;

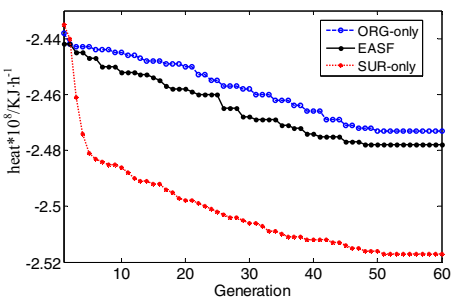
The comparative results of the convergence curve are shown in Fig. 5, and TABLE IV shows the optimization results.

TABLE III SURROGATE MODELS OF HEAT RECOVERY OPTIMIZATION OF CDU

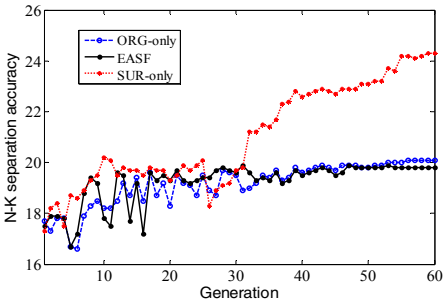
Surrogate Models	Dependent Value
Model 1	Reflux heat load
Model 2	Naphtha-Kerosene separation accuracy
Model 3	Kerosene-LGO separation accuracy
Model 4	LGO-HGO separation accuracy

TABLE IV **OPTIMAL RESULTS**

Item	Flow ratio PA1:PA2:PA3:PA4	Total heat/kJ·h^{-1}	Number of calls
Original state	0.456:0.272:0.268:0.251	2.4280×10^8	
ORG-only	0.231:0.442:0.282:0.245	2.4733×10^8	1236
SUR-only	0.788:0.822:0.990:0.932	2.5172×10^8	—
EASF	0.229:0.464:0.279:0.268	2.4782×10^8	562



a) Convergence curve for CDU



b) Best individual evolving trajectory

Fig. 5 Comparative results for CDU reflux optimization

From the optimization results it can be seen that ORG-only and EASF experiments show very similar results. Compared with the original state of CDU system, in both cases the flow ratio of overhead reflux is clearly reduced, and that of first mid-section reflux, second mid-section reflux show increase and a slight increase tendency. Due to the fact that overhead reflux has little effect on the separation accuracy of four side-cut streams and the temperatures of the streams are all low, increasing the flow rate of these two side-cut stream can improve the utilization of the waste heat for the high temperature stream. On the contrary, the flow rate of second mid-section reflux has great impact on the separation accuracy, its stream temperature appears a little higher, so it is appropriate to increase the flow rate slightly. From heat recovery point of view, compared with the single original function evaluation approach, more amount of heat is recovered by using the proposed adaptive surrogate function optimization approach as shown in Fig. 5. Further, as far as the number of original model calls are concerned, using EASF, the original model calls is only 562, compared with calls of 1236 by using ORG-only as shown in TABLE IV. It should also be noted that, it seems to get the best heat recovery result by using SUR-only as shown in Fig. 5, while a series results of pump-around flow ratios shown in TABLE IV are incredible.

To illustrate the optimization results, let's give one of the output trajectory of the surrogate models evolving with the CDU optimization process. Take Naphtha-Kerosene separation accuracy for example, Fig 5. b) is the best individual evolving trajectory for three kinds of cases, i.e., ORG-only, SUR-only, and EASF for the sake of comparison. From the figure we can see, in the early stages of the evolution, i.e. the generation g<20, EASF introduces the biggest fluctuation approximation errors, then at the middle or late stages, it showed a steady trend gradually with time. Since introduction of approximation errors is not always bad, the 'bless of uncertainty' will bring about acceleration and improving diversity in search period [26]. So, EASF's obtaining more recovery heat maybe benefit from the 'bless of uncertainty'.

V. CONCLUSIONS

To address the challenges in complex chemical process optimization, a new evolutionary optimization approach termed ASFE for enhancing surrogate modelling and optimization has been developed in this work. In contrast to the existing surrogate approaches, the method presented here uses a three layer RBF network with adaptive sampling strategy to perform data sampling and is used in evolutionary optimization, thereby effectively achieving a tradeoff between the sample numbers and optimization accuracy. Moreover, in the evolutionary process, with the improvement of the fidelity of approximate models, an adaptive individual-based model management strategy has been introduced. Therefore, the computation time for optimizing a complex system would be effectively reduced.

In this work, a new hybrid modelling framework is proposed to adaptively control the sampling quality and quantity. For the complexity of the CDU process, multiple surrogate parallel models corresponding to the optimization problem are implemented. The evolutionary optimization with the help of surrogate models is successfully applied in the CDU process operating optimization.

In the future, we are going to compare the proposed method with other modeling approaches on the basis of the thorough understanding of the optimization of the surrogate models. In addition, the adaptive modelling and surrogate management strategies will be applied to the multi-working conditions' process optimization.

ACKNOWLEDGMENT

This work is supported in part by the Natural Science Foundation of China (61503204), Natural Science Foundation of ZheJiang (LY14F030004), Science and Technology Planning Project of ZheJiang (2015C31017), the Natural Science Foundation of NingBo (2016A610092). The author gratefully acknowledges the support of K.C. Wong Magna Fund in Ningbo University.

REFERENCES

[1] M. Lluvia, O. Estopier, M. Jobson, R. Smith, "Operational optimization of crude oil distillation systems using artificial neural networks, " Computers and Chemical Engineering, vol. 59, pp.178– 185, 2013.

[2] R. Jin, W. Chen, T.W. Simpson, "Comparative studies of meta modelling techniques under multiple modelling criteria, "struct Multidisc Optim,vol. 23, pp.1–13, 2001.

[3] X. Yan, Z. Chen , X. Hu,"Chaos-genetic algorithms for optimizing the operating conditions based on RBF-PLS model," Computers & Chemical Engineering, Vol. 27, Issue 10, pp. 1393–1404, 2003.

[4] W. Lu , D. Huang, Y. Jin *et al.* "A New Strategy of Integrated Control and On-line Optimization on High-purity Distillation Process, " Chinese Journal of Chemical Engineering, Vol. 18, Issue 1, pp. 66-79, 2010.

[5] L. Liau, T. Yang, M. Tsai,"Expert system of a crude oil distillation unit for process optimization using neural networks," Expert Syst Appl, vol.2, Issue 26, pp.247-255, 2004.

[6] S. Motlaghi, F. Jalali, "Ahmadabadi M. An expert system design for a crude oil distillation column with the neural networks model and the process optimization using genetic algorithm framework," Expert Syst Appl,vol. 4,Issue 35, 2008.

[7] F. Ismail, C. Selen, "Process synthesis of biodiesel production plant using artificial neural networks as the surrogate models," Computers & Chemical Engineering, Vol. 46, 15, pp. 105–123, 2012.

[8] LM. Ochoa-Estopier, M. Jobson, R. Smith, "Retrofit of heat exchanger networks, for optimising crude oil distillation operation," Chem Eng Trans, 35, 2013.

[9] M. Lluvia, O. Estopier, M. Jobson, R. Smith, "The use of reduced models for design and optimisation of heatintegrated crude oil distillation systems," Energy, 75, pp. 5-13, 2014.

[10] M. Lluvia, O. Estopier and M. Jobson, "Optimization of Heat-Integrated Crude Oil Distillation Systems. Part I: The Distillation Model, ",Ind. Eng. Chem. Res., 54, pp.4988−5000, 2015.

[11] N. Ganesh， L. Biegler ,"A robust technique for process flowsheet optimization using simplified model approximations," Computers & Chemical Engineering,1987.

[12] I.Alhajree, G.Zahedi, Z.Manan,"Modeling and optimization of an industrial hydrocracker plant," Journal of Petroleum Science and Engineering,78,pp. 627–636, 2011.

[13] J. Caballero, I.Grossmann,"Rigorous flowsheet optimization using process simulators and surrogate models," Computer Aided Chemical Engineering, 2008 .

[14] B. Chachuata, B. Srinivasanb, D. Bonvinc,"Adaptation strategies for real-time optimization, "Computers and Chemical Engineering, 33 pp.1557–1567, 2009.

[15] W.S. Yip, T.E. Marlin, "The effect of model fidelity on real-time optimization performance, " Computers and Chemical Engineering, No. 28, pp. 267- 280, 2004.

[16] H. Yao, J. Chu. "Operational optimization of a simulated atmospheric distillation column using support vector regression models and information analysis,"Chem Eng Res Des, vol.12,90, 2012.

[17] Thacker BH, Doebling SW, Hemez FM, Anderson MC, Pepin JE, Rodriguez EA. Concepts of model verification and validation. Tech. rep. Los Alamos, USA: Los Alamos National Laboratory (LANL); 2004.

[18] H. Wang, Y. Jin et al. " Data-Driven Surrogate-Assisted Multi-Objective Evolutionary Optimization of A Trauma System," IEEE Transactions On Evolutionary Computation, 2016.

[19] Jack P.C. Kleijnen, "Kriging metamodeling in simulation: A review," European Journal of Operational Research, Vol. 192, No. 3, pp.707–716, 2009.

[20] Y. Jin, "Surrogate-assisted evolutionary computation: Recent advances and future challenges, "Swarm and Evolutionary Computation, 1,pp.61–70, 2011.

[21] Y. Jin, "A Comprehensive Survey of Fitness Approximation in Evolutionary Computation," Soft Computing, vol. 1, No. 9, pp. 3-12, 2005.

[22] Y. Jin, M. Olhofer, and B. Sendhoff, "A Framework for Evolutionary Optimization With Approximate Fitness Functions," IEEE Transactions On Evolutionary Computation,vol.6, No. 5,pp. 481- 494 , 2002.

[23] Z. Zhou, Y S. Ong, B.Prasanth,etc., "Combining Global and Local Surrogate Models to Accelerate Evolutionary Optimization," IEEE Transactions On Systems, Man, and Cybernetics-Part C: Applications and Reviews, vol. 37, No. 1, 2007.

[24] C. Sun , Y. Jin, J. Zeng, Y. Yu, "A two-layer surrogate-assisted particle swarm optimization algorithm," Soft Comput, 19, pp.1461–1475 ,2015.

[25] W. Gong ,"A Multi-Operator Search Strategy based on Cheap Surrogate Models for Evolutionary Optimization," Evolutionary Computation, Vol. 19, Issue:5 , 2015.

[26] DR Jones, "A Taxonomy of Global Optimization Methods Based on Response Surfaces, " Journal of Global Optimization, vol. 4, Issue 21, pp. 345-383,2001.

[27] D. Lim, Y. Jin, Y-S Ong, B. Sendhoff, "Generalizing surrogate assisted evolutionary computation," IEEE Transactions On Evolutionary Computation, vol. 3, No. 14, pp.329–355, 2010.