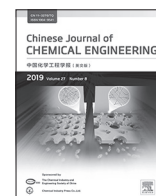




Contents lists available at ScienceDirect

Chinese Journal of Chemical Engineering

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

Article

An intelligent SVM modeling process for crude oil properties prediction based on a hybrid GA-PSO method[☆]Kexin Bi, Tong Qiu^{*}

Department of Chemical Engineering, Tsinghua University, Beijing 100084, China
 Beijing Key Laboratory of Industrial Big Data System and Application, Beijing 100084, China

ARTICLE INFO

Article history:

Received 13 September 2018
 Received in revised form 8 December 2018
 Accepted 10 December 2018
 Available online 28 December 2018

Keywords:

Intelligent properties prediction
 Support vector machine
 Hybrid GA-PSO
 TBP distillation curve fitting

ABSTRACT

Properties prediction of crude oil remains an essential issue for refineries. In this communication, an exhaustive and extendable support vector machine (SVM) intelligent prediction process has been proposed to solve this problem. A novel hybrid genetic algorithm-particle swarm optimization (GA-PSO) method was applied to optimize the SVM model. The optimization process and result demonstrated that the newly proposed GA-PSO-SVM method was more accurate and time-saving than the classical GA or PSO method. Compared with the classical Grid-search SVM, the combined GA-PSO-SVM model appeared to be more applicable for the properties prediction task. The TBP distillation curve fitting was exemplified to evaluate the performance of the developed model. The regression result demonstrated the high accuracy and efficiency of the proposed process. The model can be applied in the Industrial Internet as a plug-in, and the adaptability and flexibility is demonstrated by the implement of crude oil molecular reconstruction employing the intelligent prediction process.

© 2019 The Chemical Industry and Engineering Society of China, and Chemical Industry Press Co., Ltd.
 All rights reserved.

1. Introduction

The technology of crude oil analysis is advancing at a rapid pace nowadays, and breakthroughs are made in key properties determination [1]. For refineries, thorough evaluations of crude oil will indicate the proper selection of refining strategy, which will conduce to profit boost. However, due to the high expenditure and long time-consumption, the application of advanced analysis methods used in refineries faces many barriers. Full sets of analysis data can be obtained by labs or institutes only, and these precious datasets are not fully utilized. Machine learning research in recent years have proposed various models [2], including artificial neural network (ANN), support vector machine (SVM), etc. Potential association among crude oil properties verifies the feasibility to apply these machine learning models to do estimation of analysis result. These computer-calculated models are resource-economizing, time-saving and have a promising prospect. Visible profit will be earned if these models are successfully plugged into the Industrial Internet [3].

Among the machine learning methods, SVM is based on statistical learning theory and the structural risk minimum principle,

and performs better in small-sample learning problem with simple structure and strong generalization ability. Some studies have applied SVM as properties prediction model in crude oil analysis. A. Chamkalani [4] proposed a “two parameters” SVM model to diagnose asphaltene stability in crude oil, and the simulation result showed 0.6% average absolute relative deviation and squared correlation coefficient of 0.722. A. Kamari [5] evaluated the vaporization enthalpies of petroleum fractions with a least squares support vector machine (LSSVM) algorithm, which has an advantage in accuracy and reliability over models previously reported. A. Varamesh *et al.* [6] tried the LSSVM model for estimating the normal boiling points of crude oil fractions and demonstrated the applicability of the developed models.

Despite these fantastic characteristics, direct application of SVM in Process Systems Engineering is error-prone because of the parameters and kernel functions dependence of the model. The optimization process of SVM is a formidable task due to the strong coupling between parameters and multiple local extremum in objective function of the model. To overcome this tough problem, several optimization methodologies are tried out in implementation of SVM parameters selection. Of these algorithms, the swarm intelligence algorithm has manifested its effectiveness and precision in the field of parameter optimization, and it is suitable for SVM model training because of its robustness and rapid random-search capability. F. Kuang *et al.* [7] employed genetic algorithm

[☆] Supported by the National Natural Science Foundation of China (U1462206).

^{*} Corresponding author.

E-mail address: qiuotong@tsinghua.edu.cn (T. Qiu).

(GA) in a hybrid kernel principal component analysis (KPCA) and SVM model for intrusion detection. PJG Nieto *et al.* [8] built an SVM model with particle swarm optimization (PSO) for engine life forecasting. A high-performance optimization algorithm is conscientiously demanded in oil properties prediction for intelligent modeling.

In this work, we proposed a novel hybrid GA-PSO method to complete the optimization process. Besides, a thorough prediction model of crude oil properties is designed to make intelligent decision for refineries. Ideally, once analysis datasets of crude oil are provided, models should be obtained expeditiously and applied smoothly by refineries with properties to be predicted as output result. Generally, the properties obtained only by time-consuming experiment should be arranged as export data, and these properties are able to be regressed by easily acquired properties. As there are numerous properties to be predicted in realistic production practice, in this paper, we preliminarily focus on the true boiling point regression problem of crude oil as a function of the Engler distillation curve (ASTM D86 distillation curve), in order to verify the practicability of the proposed model based on SVM combined with swarm intelligence algorithm. A novel hybrid GA-PSO optimization method is presented as the critical procedure in intelligent modeling, and a trail of inserting the combined GA-PSO-SVM model into the molecular reconstruction process of naphtha pyrolysis is made to demonstrate that the model is adaptable and practical in industrial software [9] as a plug-in unit.

2. Methods

2.1. Intelligent prediction process as a part of Industrial Internet

A large quantity of crude oil properties prediction models can be concluded in the same category, but various requirements may be proposed by different refineries. Under such circumstances, we devised an intelligent prediction process in Fig. 1 in order to provide interchangeable solutions to these prediction problems. With the input of the collected dataset, the process will return to a packed model for refineries. During the process, the SVM model is applied to complete regression step and a hybrid GA-PSO method is developed to do parameters selection for SVM. The modeling process is auto-complete and can be utilized as a unit of Industrial Internet.

To test the capabilities of the intelligent process, the TBP distillation curve conversion problem is exemplified. Characterization of the TBP curve requires a time-consuming (48 h) and costly method, which involves a 14–18 theoretical plates rectifying column with a reflux ratio of 5:1 [10]. Several well-known empirical correlations have been developed to make conversion from simplest measured ASTM D-86 curve to TBP curve, including Riazi–Daubert Method [11], Daubert's Method [12], *etc.* Nevertheless, these correlations exhibit poor performance with tremendous estimation deviation, as was demonstrated by A. Nedelchev *et al.* [13] and D. Stratiev *et al.* [14], because the parameter-dependent

correlations cannot guarantee the flexibility, reliability and extendibility. Thus, the intelligent prediction process is applied to build a regression model with strong stability.

In addition, a molecular reconstruction process of crude oil is introduced as an application scenario of the proposed plug-in unit. The high-performance molecular reconstruction method [15] has been applied in naphtha pyrolysis software, EcSOS [9], which was developed in 2017, and the proposed SVM combined with swarm intelligence algorithm is employed to estimate the TBP curve of naphtha and the initial value which includes the distribution information in homologs in simulated annealing algorithm of molecular reconstruction.

2.2. SVM regression model

Frequently, TBP curve conversion is a small sample-learning problem, due to the inconvenient acquirement of experiment data in refineries. Thus, ε -Support Vector Regression (ε -SVR) is chosen as the regression model for its adaptability of limited amount of samples. ε -SVR is also based on VC dimension theory and structural risk minimum principle [16], making the method enhanced in generalization and global optimization performance. The classical ε -SVR can be solved by Lagrange multiplier method and Karush–Kuhn–Tucker conditions [17], and the solution is:

$$\begin{aligned} \bar{y}(\mathbf{x}) &= \sum_{\mathbf{x}_i \in SV} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}_j) + b \\ b &= \frac{1}{N_{NSV}} \left\{ \sum_{0 < \alpha_i < C} \left[y_i - \sum_{\mathbf{x}_j \in SV} (\alpha_j - \alpha_j^*) K(\mathbf{x}_j, \mathbf{x}_i) - \varepsilon \right] \right. \\ &\quad \left. + \sum_{0 < \alpha_i^* < C} \left[y_i - \sum_{\mathbf{x}_j \in SV} (\alpha_j - \alpha_j^*) K(\mathbf{x}_j, \mathbf{x}_i) + \varepsilon \right] \right\} \end{aligned} \quad (1)$$

where C and ε are parameters to be optimized, respectively represent penalty coefficient and tolerance bound; α_i and α_i^* are Lagrange multipliers; b stands for the intercept of linear regression process; \bar{y} is the output, and \mathbf{x} express the input vector; SV means the support vector; N_{NSV} means the amount of standard support vector; K donates the kernel function. In this model, Radial Basis Function (RBF) is applied as kernel function K , which can be described as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2\right) \quad (2)$$

where γ represents the distribution width in kernel function, with $\gamma = 1/2\sigma^2$, and γ also need optimization.

To summarize, three parameters (C , γ , and ε) need to be optimized in this model. The values of these parameters will determine the learning ability and generalization ability.

2.3. Hybrid GA-PSO optimization method for SVM parameters selection

SVM training process equals to the parameters selection to achieve the least prediction error. The root-mean-square error

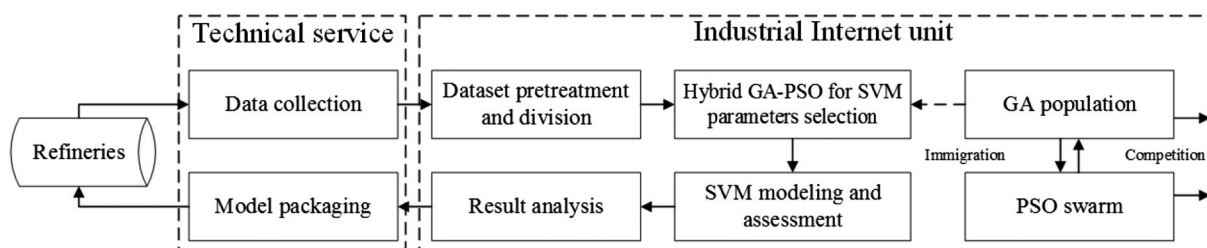


Fig. 1. The intelligent process for crude oil properties prediction.

(RMSE) is applied for precision assessment. As previously mentioned, the parameters to be optimized are C , ε and γ . These parameters should be restricted in order to guarantee the learning capacity and generalization capacity of the model. The optimization problem can be described as:

$$\begin{aligned} \max \quad & -RMSE = -\sqrt{\frac{1}{n} \sum (\bar{y}_i - y_i)^2} \\ \text{s.t.} \quad & \log_2 C \in [lb_1 ub_1], \log_2 \gamma \in [lb_2 ub_2], \log_2 \varepsilon \in [lb_3 ub_3] \end{aligned} \quad (3)$$

where \bar{y}_i and y_i represent the predicted and actual output of the model. Lower bound (lb_i) and upper bound (ub_i) can be assigned according to model requirements.

Swarm intelligence algorithms are widely used in simulation to optimize multiple local extremum problem [18]. There into, GA draws on the experience of natural selection and Mendel's genetic law, and runs computing process with distinctive chromosome coding and decoding operations [19]. The combination of SVM training and GA method can be executed as the following steps. First, implement population initialization by chromosome coding function. Second, do SVM training for each initial individual, and record the training error (RMSE) as fitness function value. Then update the genotype of each individual by selection, crossover and mutation every loop. Finally, when termination condition is tenable, decode the individual with best fitness as output result.

PSO is another swarm intelligence algorithm with the simulation of social behavior such as birds flocking [20,21]. Unlike GA method, PSO is established with a simpler architecture, and the application in SVM parameter optimization is as follows. First, perform a random generation process of population (called swarm). The dimension of each individual (called particle) equals to the amount of optimization parameters, and the vectors in the initial swarm are used to calculate the fitness (RMSE). Then, update the velocity and position of particles with searching functions and calculate the fitness again, until the termination criterion is satisfied. The velocity of particles can be determined by the historical extremum of each particle and swarm.

When applying in SVM parameter optimization, GA and PSO are not able to complete the formidable task. For GA method, the local-searching capacity is outstanding, but the coding and decoding process increases the complexity of calculation, and the iteration steps are added because the feedback information of loop network is unused [19]. The optimization process is time-consuming and sensitive to population initialization. For PSO method, the calculation speed of each step is accelerated because of the lower algorithm complexity. However, a local optimal solution is sometimes obtained. Besides, the poor local-searching performance of PSO could make the convergence slow down at the later stage and produce an inaccurate result [20,21], so a new optimization method needs to be developed sufficiently.

Taking advantage of both GA and PSO, we proposed a novel hybrid GA-PSO optimization method for SVM parameter selection. Two divided population are initialized in the beginning, and separately evolved in iteration process. The fundamental frame of two routes is remained, and some modifications are added into every iteration step to accelerate the convergence rate and prevent premature convergence, which include uniform population initialization, filtering out similar individuals and applying self-adaptive parameter adjusting strategy. (See Fig. 2.)

The key steps in this new method are based on the cooperation and competition between populations. In each generation, PSO and GA are in implementation simultaneously. Besides, competitive particles (with lower RMSE) in PSO population are imported in GA population and eliminate the worse individuals (with higher RMSE) to remedy the lack of feedback in GA method, and the better individuals (with lower RMSE) in GA population are imported in PSO population to add another guidance for searching. To maintain the amount of population, inferior individuals are altered by the same amount of better ones. The supplementary feedback of GA method surmounts the problem of oversensitivity to population initialization and accelerates the convergence in starting generations. The emigrant individuals of GA to PSO make up the deficiencies that the optimal solution or inaccurate solution might be obtained because of poor local-searching performance in PSO method.

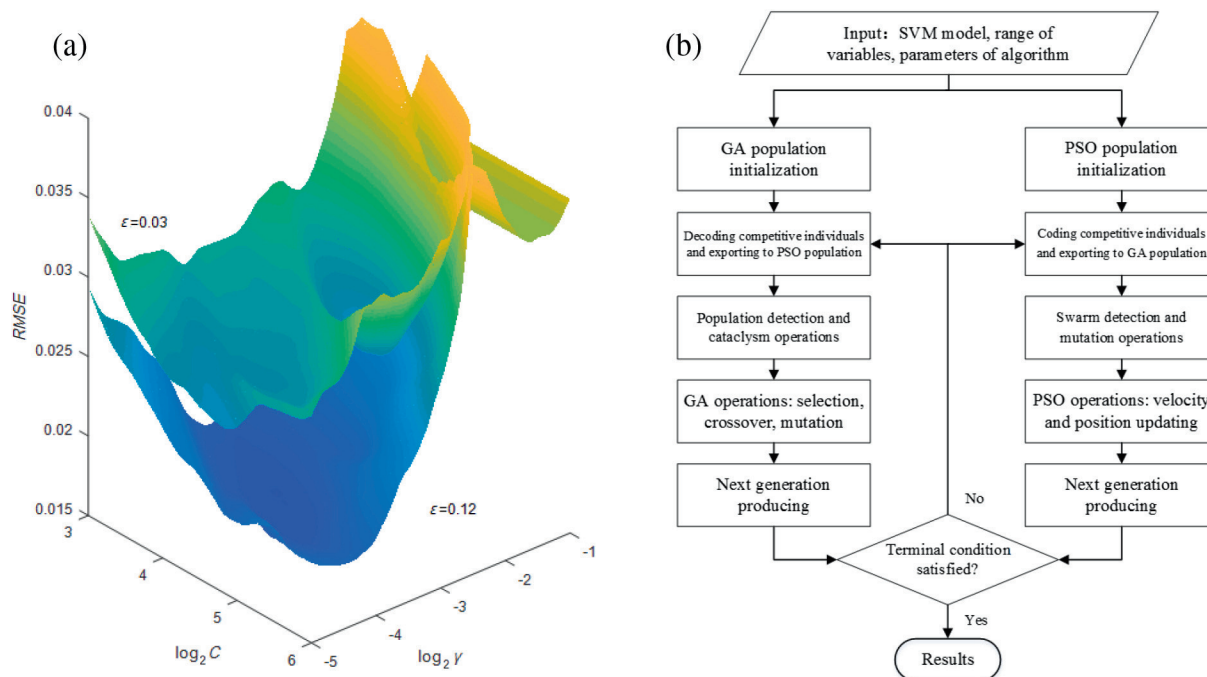


Fig. 2. (a) The RMSE graph of SVM model as a function of C , γ and ε . (b) The structure diagram of hybrid GA-PSO method.

3. Results and Discussion

3.1. TBP distillation curve regression with intelligent prediction process

The experimental TBP data and ASTM D-86 data were collected from Chinese refineries. One hundred seventy-three groups of boiling point data for naphtha have been treated by the intelligence prediction model in Fig. 1. One hundred thirty-five sets of data (about 80% of the whole dataset) have been divided into the training set and carried out a 5-fold cross validation. The rest 38 sets of data were remained as test set. The division method was based on the 'train_test_split' algorithm from sklearn function library [22]. The SVM regression result can be concluded as:

$$\overline{TBP}_{v_k} = \sum_{T_i \in SV} (\alpha_i - \alpha_i^*) K(T_i T_j) + b \quad (4)$$

$$T_i, T_j = (D86_{10\%}, D86_{30\%}, D86_{50\%}, D86_{70\%}, D86_{90\%})$$

where \overline{TBP}_{v_k} and $D86_{v_k}$ represents the predicted temperature points on TBP distillation curve and actual temperature points on ASTM D-86 distillation curve, with $v_k = 10\%, 30\%, 50\%, 70\%, 90\%$.

After the consolidation and scaling pre-treatment of original dataset, the hybrid GA-PSO method should be operated for SVM parameters selection. The pivotal settings in GA method can be

referred to the adaptive genetic algorithm proposed by M. Srinivas *et al.* [23], and settings in PSO method can be referred to the self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients proposed by A. Ratnaweera *et al.* [24]. The uniform population initialization was introduced by uniform distribution in variable space. The population filtering and parameters self-adaption method were referred to the literature mentioned above. To compare the new GA-PSO method with the isolated GA or PSO method, the same SVM parameters optimization problems were also simulated by the GA or PSO with same control coefficients. The optimization programs were simulated by Matlab R2017a with Inter Core i7-7700HQ CPU at 2.80GHz. The convergence trend of each method is plotted in Fig. 3.

As shown in Fig. 3, in hybrid method, fitness of both GA and PSO converged at 0.0576 (10th generation) or 0.0575 (16th generation), due to the cooperation and competition mode. When these methods are independently applied in the optimization process, the convergence speed was decreased with isolated GA in 33rd generation (at 0.0577) and isolated PSO in 31st generation (at 0.0596). Indeed, the hybrid method not only enhanced the accuracy of both GA and PSO, but also accelerated the convergence. The evolutionary trend of hybrid method (in the circumstance of $\varepsilon = 0.125$) is shown in Fig. 4. The grid search optimization was applied to draw contour graph.

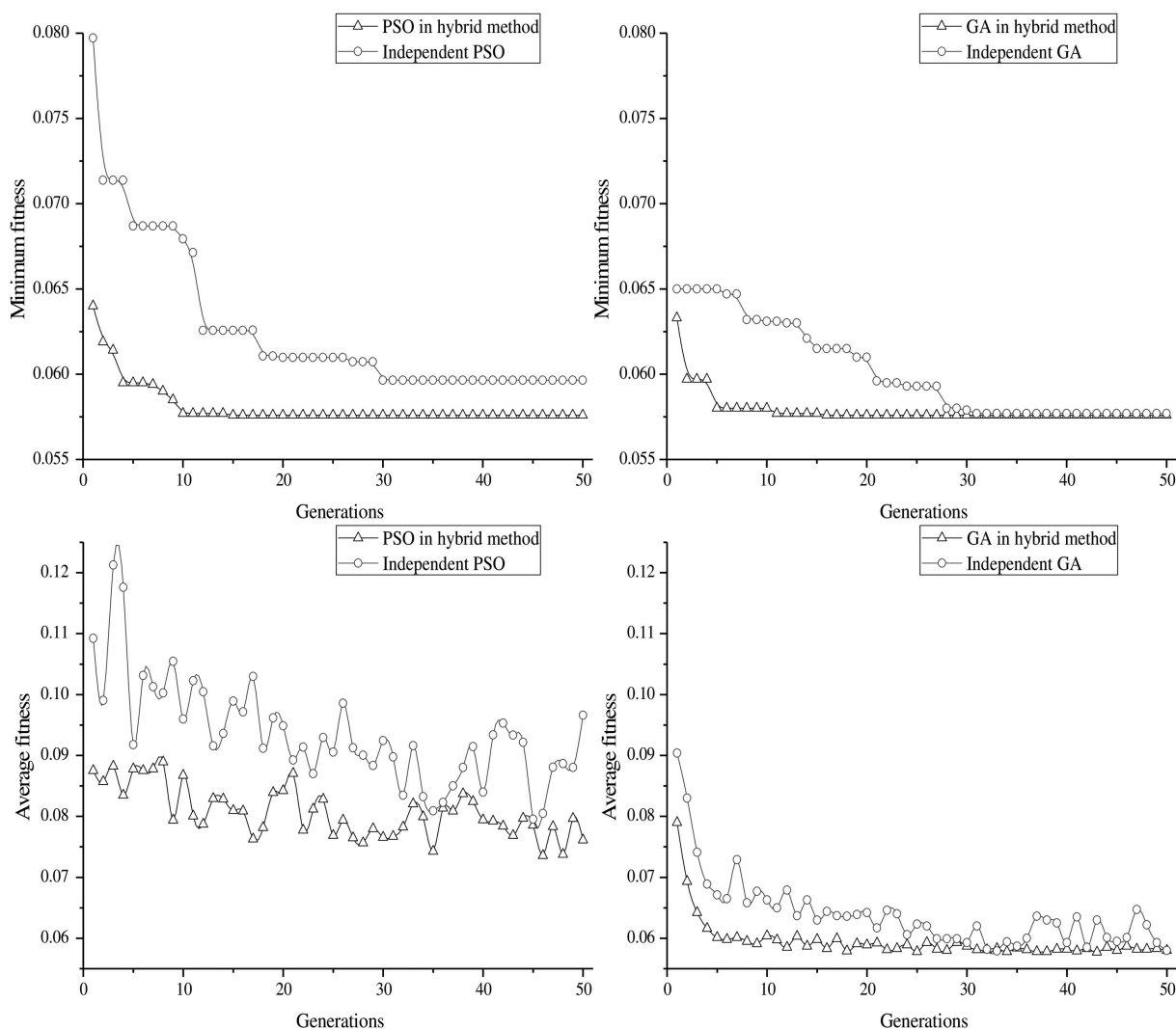


Fig. 3. Convergence trend of each method.

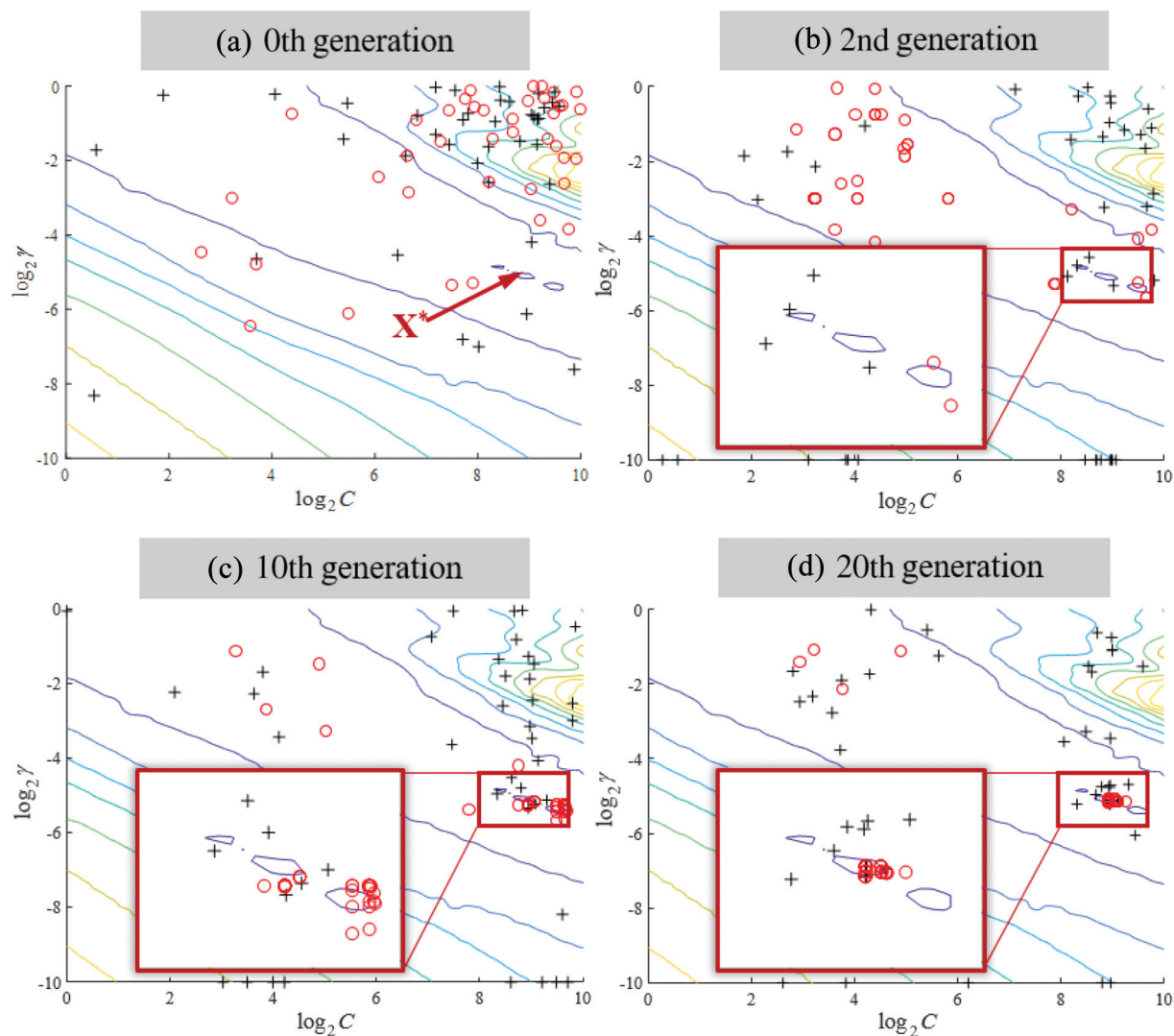


Fig. 4. The evolutionary trend of hybrid method in contour graph when $\varepsilon = 0.125$, + for PSO, o for GA.

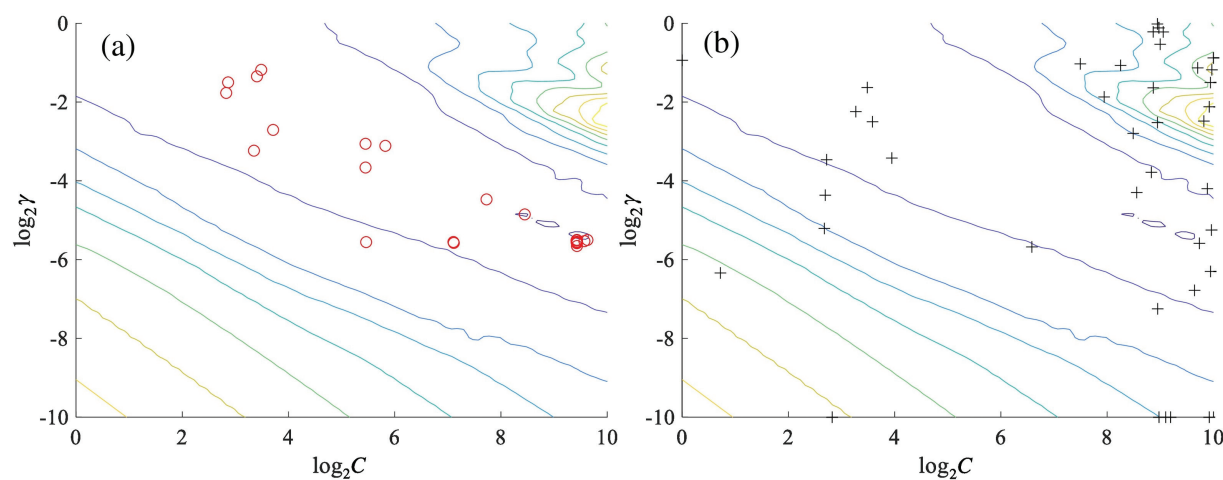


Fig. 5. The evolutionary trend of independent method (20th generation) in contour graph when $\varepsilon = 0.125$, + for PSO, o for GA.

As shown in Fig. 4 (a), the global optimal solution area was labeled with x^* , and the initial populations of GA and PSO were generated uniformly and randomly on the Cartesian coordinate

(so dots seemed to distribute non-uniform in logarithmic coordinate). After earliest searching and information exchange in two populations, both methods found the local optimal solutions in

Table 1
Detailed evaluation results of different methods

Method	Whether global optimal solution is find	Average RMSE on training set	Average RMSE on test set	Time consumption of parameters selection/s	Time consumption of other steps/s
PSO-SVM	No	0.0596	0.0714	5.54	13.11
GA-SVM	Nearly in 33rd generation	0.0577	0.0664	36.46	11.59
GA-PSO-SVM	Nearly in 10th generation	0.0575	0.0620	10.59	12.88
Grid search-SVM	No	0.0598	0.0656	490.6	12.76

2nd generation. However, in 10th generation, a large proportion of individuals in GA method tended to the local optimal area on the right side of global optimal solution. Simultaneously, particles in PSO kept the global optimal solution and shared the information with GA population by immigration of competitive particles. GA population also transmitted the local-searching information by export competitive individuals to PSO swarm. In 20th generation, both populations converged to global optimal solution.

When PSO or GA runs independently to solve this problem, the optimization result might be undesirable. As shown in Fig. 5, in the 20th generation, isolated GA population (a) was lost in local optimal solution, and isolated PSO population (b) felt it hard to complete local searching in limited steps. The lack of cooperation and competition between two methods reduced the accuracy and reliability of optimization process.

To evaluate the performance of the intelligent prediction model based on different optimization method, the whole process was carried out to do a three-parameter selection for SVM model. The specific evaluation indicators and results are listed in Table 1. Among these methods, PSO-SVM completes the parameters selection first, but the solution is undesirable. It might because the randomness of PSO is too strong to do local searching. Grid search method is able to find the global optimal solution in theory, but the time consumption will be inestimable if the step length is too short. In this problem, GA-SVM and GA-PSO-SVM both found the global optimal solution nearly. However, the hybrid optimization method apparently had it over the GA method considering the calculation time. The lack of feedback in GA method made the fitness value drop slowly, and the reliability of the GA solution is inferior because the GA method is too sensitive to the population initialization. The adjunction of PSO into GA solves these problems and guides the GA population to the global optimal solution within a rather transient evolution process. The hybrid GA-PSO method finds the solution only in the 10th generation, and it appears to be time-saving. The time-consumption of other steps in four methods is approximate, so the efficiency of optimization step

Table 2
Detailed performance of GA-PSO-SVM method versus conventional method on a single case

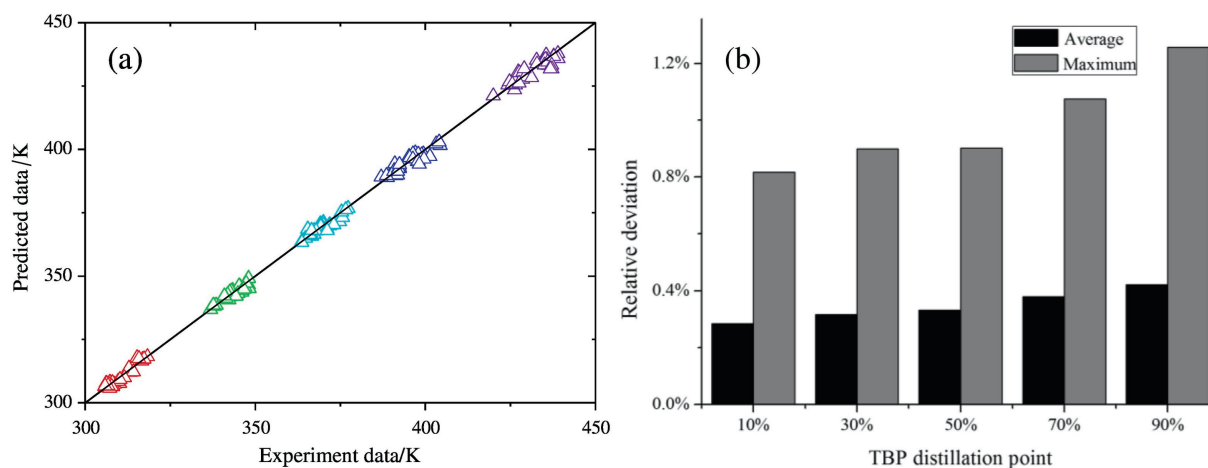
Method	Average RMSE of TBP distillation data	Average RMSE of carbon number distribution data	Average deviation of predicted mass percentage/wt%
GA-PSO-SVM	0.0613	0.0975	0.5115
Conventional	0.1150	0.1613	0.7469

determines the performance of the whole process. Overall, the intelligent model based on hybrid GA-PSO method had the better performance than other methods mentioned in this paper.

The test set prediction error of the proposed GA-PSO-SVM process is shown in Fig. 6. Five groups of distillation temperature point predicted results (TBP_{10%}, TBP_{30%}, TBP_{50%}, TBP_{70%}, TBP_{90%}) are contrasted to the experiment data, and the absolute relative deviation distributions in Fig. 6 (a) illustrates that the process has low scatter around zero. By calculating the relative deviation on test set, we found that the average relative deviation is below 0.5%, and the max is below 1.3%.

3.2. Intelligent prediction model in molecular reconstruction process

A case study based on the molecular reconstruction process was exemplified here to make an effort to plug the intelligent prediction program into the EcSOS software [9]. According to the high-performance molecular reconstruction method [15], the TBP distillation curve and the gamma distribution parameter of carbon number in homologs are needed before the whole reconstruction process, so we applied the proposed GA-PSO-SVM method to regress these values and made a comparison with results of the conventional correlations method (Riazi–Daubert method for TBP regression, PCA regression for distribution parameter regression) using the same input data utilized in the high-performance molecular reconstruction model. The detailed result is shown in Table 2 and Fig. 7.

**Fig. 6.** Comparison between predicted data and experiment data. (a) Scatter diagram (b) relative deviation analysis.

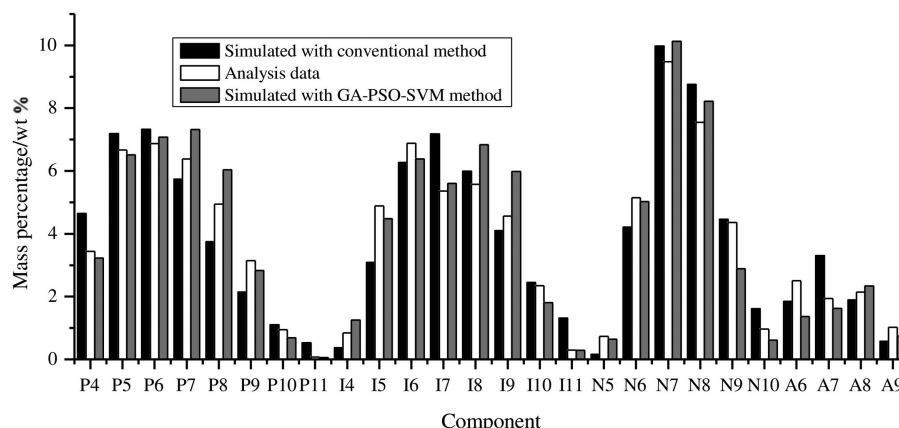


Fig. 7. Molecular reconstruction result of naphtha using different method. The name of component was composed of their homologous group (*n*-Paraffin, *Iso*-paraffin, Olefin, Naphthene, and Aromatic) and the carbon number.

As is shown in Table 2 and Fig. 7, the molecular reconstruction model in Industrial Internet showed a better performance with the plug-in of GA-PSO-SVM method. The average deviation and time consumption of the model was declined because a more accurate estimation result of TBP distillation curve and carbon number distribution was provided. The proposed intelligent prediction process was proved to be a feasible plug-in in the Industrial Internet.

4. Conclusions

There are several approaches to complete the intelligence TBP curve prediction. SVM is applied as regression model, because the global optimal solution is easier to find in the SVM convex optimization process, and overfitting risk is avoided with soft-margin treatment. However, parameter selection of SVM model is still formidable to solve, and classical grid search method is indeed time-consuming in local-searching process because too much data point need to be calculated. The newly developed hybrid GA-PSO for SVM parameter optimization provides a feasible solution to this conundrum. The overall accuracy of the GA-PSO-SVM surpasses all models mentioned in this paper, and the time-consumption is acceptable for industrial application requirement. The introduction of cooperation and competition mechanism makes the populations in hybrid GA-PSO method tend to the global optimal solution faster, and the populations in single GA and PSO method do not appear to converge in limited hereditary algebra.

Moreover, the flexibility and extendibility of the proposed prediction process guarantee the feasible application in other prediction problems of crude oil, and the whole process is demonstrated to be practical as a unit in the pyrolysis software, and further in the Industrial Internet. Besides, an intelligent 'data collection-model packaging' loop is expected to help the refineries estimate the crude oil thoroughly and be more profitable in refining process. The properties prediction process can be inserted into the Industrial Internet to create value for petrochemical enterprises.

References

- [1] J.C. Chen, Z. Fang, T. Qiu, Molecular reconstruction model based on structure oriented lumping and group contribution methods, *Chin. J. Chem. Eng.* 26 (8) (2018) 1677–1683.
- [2] S.B. Kotsiantis, Supervised machine learning: a review of classification techniques, *Informatica* 31 (2007) 249–268.
- [3] P. White, The power of the industrial internet: turning data into insight and action, *J. Pet. Technol.* 66 (11) (2014) 90–93.
- [4] A. Chamkalani, A.H. Mohammadi, A. Eslamimanesh, et al., Diagnosis of asphaltene stability in crude oil through "two parameters" SVM model, *Chem. Eng. Sci.* 81 (2012) 202–208.
- [5] A. Kamari, M. Sattari, A.H. Mohammadi, et al., Modeling of the vaporization enthalpies of petroleum fractions, *Fluid Phase Equilib.* 412 (2016) 228–234.
- [6] A. Varamesh, A. Hemmati-Sarapardeh, B. Dabir, et al., Development of robust generalized models for estimating the normal boiling points of pure chemical compounds, *J. Mol. Liq.* 242 (2017) 59–69.
- [7] Fangjun Kuang, Weihong Xu, Siyang Zhang, A novel hybrid KPCA and SVM with GA model for intrusion detection, *Appl. Soft Comput.* 18 (2014) 178–184.
- [8] P.J. García Nieto et al., Hybrid PSO-SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability, *Reliab. Eng. Syst. Saf.* 138 (2015) 219–231.
- [9] Zhou Fang, Tong Qiu, Weiguo Zhou, Coupled simulation of recirculation zonal firebox model and detailed kinetic reactor model in an industrial ethylene cracking furnace, *Chin. J. Chem. Eng.* 25 (8) (2017) 1091–1100.
- [10] A. Ba, E. Eckert, T. Vaněk, Procedures for the selection of real components to characterize, *Chem. Pap.* 57 (1) (2003) 53–62.
- [11] Mohammad R. Riaz, Thomas E. Daubert, Analytical correlations interconvert distillation-curve types, *Oil Gas J.* 84 (34) (1986) 51–57.
- [12] T.E. Daubert, Petroleum fraction distillation interconversions, *Hydrocarb. Process.* 73 (1994) 9.
- [13] Angel Nedelchev et al., Boiling point distribution of crude oils based on TBP and ASTM D-86 distillation data, *Pet. Coal* 53 (4) (2011) 275–290.
- [14] D.S. Stratiev et al., Evaluation of approaches for conversion of ASTM into TBP distillation data of oil fractions, *Oil Gas Eur. Mag.* 4 (2014) 216–221.
- [15] Kexin Bi, Tong Qiu, A high-performance molecular reconstruction method with parameter initialization based on PCA, *Computer Aided Chemical Engineering*, vol. 44, Elsevier, 2018, pp. 2005–2010.
- [16] Wang Jianlin, Xuying Feng, Tao Yu, A geometric approach to support vector regression and its application to fermentation process fast modeling, *Chin. J. Chem. Eng.* 20 (4) (2012) 715–722.
- [17] L.I.U. Han, Liu Ding, Research on natural gas short-term load forecasting based on support vector regression, in: 5th World Congress on Intelligent Control and Automation, IEEE, 2004, <https://doi.org/10.1109/WCICA.2004.1343263>.
- [18] Yuanbin Mo, Dezhao Chen, Shangxu Hu, Particle swarm optimization for multi-objective process system optimization problems, *J. Chem. Eng. Chin. Univ.* 22 (1) (2008) 94.
- [19] Manoj Kumar et al., Genetic algorithm: Review and application, *Int. J. Inf. Technol. Knowl. Manag.* 2 (2) (2010) 451–454.
- [20] Alec Banks, Jonathan Vincent, Chukwudi Anyakoha, A review of particle swarm optimization. Part I: background and development, *Nat. Comput.* 6 (4) (2007) 467–484.
- [21] Alec Banks, Jonathan Vincent, Chukwudi Anyakoha, A review of particle swarm optimization. Part II: hybridisation, combinatorial, multicriteria and constrained optimization, and indicative applications, *Nat. Comput.* 7 (1) (2008) 109–124.
- [22] Fabian Pedregosa et al., Scikit-learn: machine learning in python, *J. Mach. Learn. Res.* 12 (Oct 2011) 2825–2830.
- [23] Mandavilli Srinivas, Lalit M. Patnaik, Adaptive probabilities of crossover and mutation in genetic algorithms, *IEEE Trans. Syst. Man Cybern.* 24 (4) (1994) 656–667.
- [24] Asanga Ratnaweera, Saman K. Halgamuge, Harry C. Watson, Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients, *IEEE Trans. Evol. Comput.* 8 (3) (2004) 240–255.