

# Prediction of crude oil refractive index through optimized support vector regression: a competition between optimization techniques

Amin Gholami<sup>1</sup> · Hamid Reza Ansari<sup>2</sup> · Seyedasad Hosseini<sup>3</sup>

Received: 13 September 2015 / Accepted: 17 January 2016 / Published online: 1 February 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** Refractive index (RI) provides valuable information about various reservoir engineering calculations, making it a key parameter for characterizing crude oils. Determination of this index through experiment is capital-intensive, time consuming, and also toil. Hence, it is essential to search for an efficient and accurate estimation of crude oil RI. In this study, an intelligent approach, based on optimized support vector regression (SVR), is introduced to find a quantitative correlation between crude oil RI and SARA (saturate, aromatic, resin, and asphaltene) fraction data. Optimization of SVR is implemented through three searching approaches, viz. hybrid of grid and pattern search (HGP), genetic algorithm (GA), and imperialist competitive algorithm (ICA). Using these approaches, three models are constructed and tested on experimental data gathered from open source literature. To evaluate the performance of these models, their outputs are compared with corresponding experimental data in terms of statistical criteria. The comparative study clearly shows the advantage of ICA over its rivals (GA and HGP) in optimizing the SVR parameters. ICA optimized support vector regression results in an  $R^2$  of 0.9971 and MSE of  $1.48548e-05$  demonstrating its efficacy in obtaining crude oil refractive index from SARA data.

**Keywords** Refractive index (RI) · Support vector regression (SVR) · Hybrid grid and pattern search (HGP) · Genetic algorithm (GA) · Imperialist competitive algorithm (ICA)

## Introduction

Refractive index (RI) is one of foremost representative properties of crude oils. This optical parameter is a fundamental physical property pertaining to the speed of light in crude oil (Chamkalani et al. 2014). Recently, this parameter has achieved growing attention in petroleum industry, mainly because of its application in various calculations regarding crude oil specific compositions (El Ghandoor et al. 2003). Touba et al. (1997) investigated the applicability of crude oil RI for defining various reservoir fluid characterizing properties. Their results demonstrated that crude oil RI can be used as a suitable property for predicting other crude oil properties including those controlling PVT behavior of hydrocarbon and surface tension. Taylor et al. demonstrated that the onset of asphaltene precipitation could be detected from RI measurement (Taylor et al. 2001). Fan et al. (2002) examined the relation between crude oil RI and asphaltene stability in crude oil. They suggested that the difference between crude oil RI and crude oil RI at the onset of asphaltene precipitation could be employed as a decisive factor for the assessment of asphaltene stability in crude oils (Fan et al. 2002). Buckley and Wang (2002) showed that the wettability altering tendency is related to crude oil RI. Riazi and Al-Otaibi (2001) attempted to shed a new light on the relation between the viscosity and RI. They proposed a simple mathematical-based approach which can estimate the viscosity of crude oils from their RI data.

✉ Amin Gholami  
amingholamiput@yahoo.com; AGHolami@iooc.co.ir

<sup>1</sup> Reservoir Engineering Division, Iranian Offshore Oil Company, Tehran, Iran

<sup>2</sup> Abadan Faculty of Petroleum Engineering, Petroleum University of Technology, Abadan, Iran

<sup>3</sup> Department of Petroleum Engineering, Sharif University of Technology, Tehran, Iran

Considering the importance of crude oil RI, introduction of a reliable technique for accurate determination of this parameter is valuable. Conventional refractometers have long been employed as the leading method of direct measurement for determination of crude oil RI. However, the conventional refractometer is incapable of measuring the RI of heavy crude oils, mainly due to its opacity (Taylor et al. 2001). Recently, a novel RI measurement technique, the so-called capillary tube interferometer, has been introduced by El Ghandoor et al. (2003) for laboratory determination of heavy crude oil RI. This method can determine the RI of heavy crude oil, as well as light crude oil, with high precision. Although aforementioned experimental technique measures the RI of crude oil with high accuracy, utilizing this method for quantifying crude oil RI is associated with some difficulties as experimental implementation is costly, time consuming, and also labor intensive. Due to drawbacks encountered in various methods of experimental measurement, development of a crude oil RI model could facilitate exact calculation of crude oil RI from readily available experimental data. Several researchers have tried to develop a predictive model for determining the RI (Chamkalani et al. 2014; Fan et al. 2002; Buckley and Wang 2002; Chamkalani et al. 2012; Chamkalani 2012; Gholami et al. 2014a, b; Vargas and Chapman 2010; Zargar et al. 2015; Tatar et al. 2015). The published models available in the literature for estimating the value of crude oil RI can be divided into two main categories. In the first group of models, RI of crude oils is estimated using their density. Buckley and Wang (2002) developed a correlation relating crude oil RI to its density through linear regression. Chamkalani (2012) also presented a similar correlation to that of Buckley and Wang (2002). Vargas and Chapman (2010) evaluated the applicability of the One-Third rule in hydrocarbon and crude oil systems. They introduced a novel model, so-called Lorentz–Lorenz expansion, for building a correlation between RI and density of crude oils. In the second group of models, the RI of crude oil is calculated using SARA (saturate, aromatic, resin, and asphaltene) fraction data. The concept of the mutual dependence of crude oil RI and SARA fraction data was introduced by Fan et al. (2002). They developed an empirical correlation between the SARA fraction data and RI of crude oil. Chamkalani (2012) established another empirical correlation to estimate the value of crude oils RI. Although aforementioned empirical correlations are valuable, they are faced with some shortcomings, the main one being unsatisfactory accuracy.

Recently, intelligence based models have emerged as powerful techniques of solving complex problems encountered in petroleum industry. These models have achieved encouraging results in modeling phenomena and parameters such as asphaltene precipitation, minimum

miscible pressure, viscosity, porosity, permeability, bubble point pressure (Asoodeh et al. 2014a, b; Bagheripour et al. 2015; Gholami et al. 2014c; Rasuli Nokandeh et al. 2012; Hemmati-Sarapardeh et al. 2013; Afshar et al. 2014; Gholami et al. 2015; Naseri et al. 2014; Shokoo Saljooghi and Hezarkhani 2014, 2015; Ashoori et al. 2010). Moreover, intelligent models have been used for developing a quantitative correlation between the RI and SARA fraction data. Chamkalani et al. (2012) proposed a robust mathematical approach based on the least square support vector machine for prediction of crude oil RI from SARA fraction data. Later, Chamkalani et al. (2014) improved their approach by combining two optimization tools, viz. coupled simulated annealing (CSA) and Nelder and mead simplex method (NM). Gholami et al. (2014a) presented a new mathematical approach based on fuzzy logic to predict the value of crude oil RI from SARA fraction data. In another report, Gholami et al. (2014b) attempted to create a correlation between these two parameters. They published a smart model based on neural network optimized by a genetic algorithm pattern that searches for estimation of crude oil RI as a function of SARA fraction data. Zargar et al. (2015) included particle swarm optimization in fuzzy formulation and employed construction model for estimation of refractive index of crude oil. Crude oil RI obtained from above empirical correlations and intelligent models are employed as a criterion for diagnosis of asphaltene stability in crude oil. Although aforesaid intelligence models can quantify the crude oil RI with reasonable accuracy, there exists a quest for introduction of more reliable algorithms.

Current work is aimed to offer an innovative algorithm based on optimized SVR for creating a quantitative formulation between the RI of crude oil and SARA fraction data. Optimization approaches used in this study for optimizing SVR are HGP, GA, and ICA. Optimization techniques were embedded in SVR structure for computing the optimal value of parameters found in SVR formulation. Also, to identify which optimization approach has the best performance, the comparison study based on statistical criteria is implemented. It is observed that coupling ICA with SVR achieves better performance in estimating crude oil RI.

## Model description

### Support vector regression

Support vector regression (SVR) is a powerful approximation technique based on statistical learning theory proposed by Vapnik (1995). In this technique, the input data,  $x$ , are mapped from original space into a multidimensional

feature space using a kernel function and then, a linear fitting function in this space is computed. The general SVR estimating function is a linear function as (Vapnik 1998):

$$f(x) = \sum_{i=1}^l (a_i - a_i^*)K(x_i, x) + b \tag{1}$$

where  $b$  is the offset regression estimation,  $a_i$  and  $a_i^*$  are Lagrange multipliers and  $K(x_i, x)$  is the kernel function. The Lagrange multipliers are obtained by minimizing a regularized risk function as follows (Steinwart and Christmann 2008; Vapnik 1982, 1998):

$$R_{\text{reg}}[f] = \frac{1}{2} \|\omega\|^2 + C \sum_{i=0}^n L(f(x_i) - y_i) \tag{2}$$

where  $\|\omega\|^2$  is characterized as the model complexity. The term  $C$  is a constant that controls the tradeoff between training error minimization and model complexity. On the right hand side of Eq. (2),  $n$  refers to the size of the training data,  $y$  is the target data and  $L(f(x_i) - y_i)$  is the  $\epsilon$ -insensitive loss function defined by:

$$L(f(x) - y) = \begin{cases} 0 & \text{if } |y - f(x)| \leq \epsilon \\ |y - f(x)| - \epsilon & \text{Otherwise} \end{cases} \tag{3}$$

The  $\epsilon$ -insensitive loss function can be visualized as a tube that surrounds the training data with a radius equal to  $\epsilon$ . The parameters  $C$  and  $\epsilon$  can be selected by the optimization algorithm. Another key parameter in SVR is the kernel parameter. In this study, a Gaussian radial basis function (RBF) is used as the kernel function and  $\gamma$  is its main parameter. The feature space of the RBF kernel is Hilbert space or infinite dimensional space. The RBF kernel is defined as (Vert et al. 2004):

$$K(x_i, x) = \exp(-\gamma \|x_i - x\|^2) \tag{4}$$

A more detailed account of SVR method can be found in other papers and reviews (Vapnik 1982, 1995, 1998; Al-Anazi and Gates 2010). Performance of SVR methodology is heavily dependent on the determination of parameters included in its formulation (Al-Anazi and Gates 2010; Ansari and Gholami 2015a, b). The following sections introduce the searching tools employed in this study for achieving the optimal values of SVR parameters.

### Hybrid of grid and pattern search

Hybrid of grid and pattern search (HGP) is a conventional optimization method employed for attaining accurate values of SVR parameters. This methodology is the combination of grid search and pattern search techniques. The grid search is implemented using values of parameters across a specified given search range by geometric steps. The pattern search commences at the mid-point of the

given search range and takes trial steps in each direction for each parameter. If the performance of the model is improved, the search mid-point shifts to the new point and the process is iterated. If no improvement in the model fit is achieved, the step size is reduced and the search is tried again. The pattern search terminates when the search step size is less than a pre-specified tolerance. Main shortcoming of the grid search method is its computationally expensiveness, due to the need for assessing the model at many points within the grid for each parameter. A pattern search, on the other hand, generally requires far fewer evaluations of the model than a grid search. However, it may converge to a local in lieu of global optimal point. Therefore, for obtaining an efficient optimization tool, both methods should be combined in a systematic way. In this study, optimization starts with grid search trying to find a region close to the global optimum point. Subsequently, a pattern search is conducted over the narrowed search range surrounding the best point found by the grid search. This hybrid method improves the searching methodology for determining optimal values of SVR parameters with high accuracy (Al-Anazi and Gates 2012).

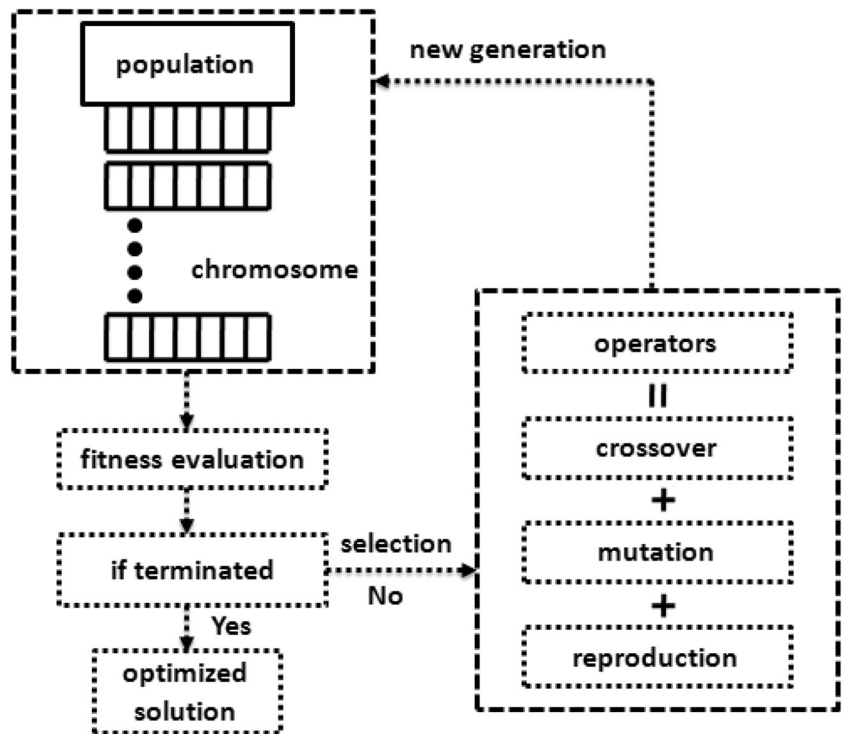
### Genetic algorithm

Genetic algorithm (GA), first introduced by Holland, is a global optimization strategy based on mathematical simulation of biological evolution (Holland 1975). In the GA terminology, potential solutions are called chromosomes. Optimization is started with initial population of random chromosomes. Evolution of each chromosome is achieved through a fitness function introduced in the problem. In each generation (algorithm iteration), the new solution is obtained through three main operations: selection, crossover, and mutation. The selection operator chooses those parents that contribute to the population in the next generation. The crossover produces children from parents, and mutation alters some gene values in a chromosome, randomly. The aforementioned process is repeated until the algorithm converges to global optima. The flowchart of GA is illustrated in Fig. 1. More details about GA can be found in Mohaghegh (2000), Khishvand and Khomehchi (2012) and Assareh et al. (2010).

### Imperialist competitive algorithm

Imperialist competitive algorithm (ICA) can be considered as a social counterpart of the GA. The ICA was introduced first by Atashpaz-Gargari and Lucas based on computer simulation of human social evolution in order to solve optimization problems (Atashpaz-Gargari and Lucas 2007). This optimization strategy has revealed great performance in the achievement of global optima (Atashpaz-Gargari et al. 2008; Alikhani-koupaei and Abdechiri 2010;

Fig. 1 General flowchart of GA



Ansari 2014; Khabbazi et al. 2009; Rajabioun et al. 2008). The ICA utilizes a special terminology, i.e., the counterpart of chromosomes in the GA is called countries in the ICA; the function expected to be optimized is called cost function; each iteration of algorithm is called a decade. In this method, based on cost values, some of the best (least values) initial countries are selected as imperialists and other countries are called colonies. Each imperialist takes control of some colonies and they constitute an empire.

Similar to the GA, ICA uses three main operations during the optimization process: assimilation, revolution, and competition. Assimilation is the process of moving the colonies toward the imperialist along different socio-political axis such as culture, language, religion, etc. (Fig. 2). In the Fig. 2,  $d$  is the distance between the colony and the imperialist state and  $x$  is a random variable with uniform (or any proper) distribution which can be obtained as (Atashpaz-Gargari and Lucas 2007):

$$x \sim U(0, \beta \times d) \tag{5}$$

where  $\beta$  is a constant number greater than one. Revolution causes the position of some countries change along social axis in a relatively short period time. The revolution increases the ability of the solution search in the algorithm and prevents early convergence of the algorithm to local minimums. During assimilation and revolution, a colony might reach to a position with lower cost than the imperialist and thus, the positions of the imperialist and the colony will be changed in this case. Furthermore, all

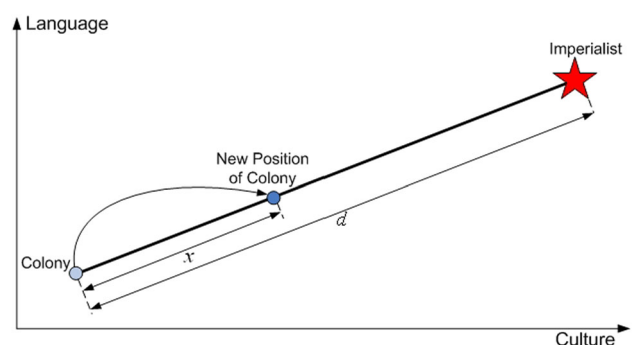


Fig. 2 Movement of colonies toward their relevant imperialist along the language and culture axis (Atashpaz-Gargari and Lucas 2007)

empires try to take the possession of colonies in other empires and control them. In this case an imperialist competition will happen. This process is done based on total cost (T.C) of the empires which can be obtained by:

$$TC_n = Cost(imperialist_n) + \xi \text{ mean}\{Cost(colonies \text{ of } empire_n)\} \tag{6}$$

where  $\xi$  is a coefficient which controls the percentage of contribution of the colonies costs in the total cost. The normalized total cost (NTC) is obtained by:

$$NTC_n = TC_n - \max\{TC_i\} \tag{7}$$

Then, the possession probability (pp) of each empire is given by:

$$pp_n = \left| \frac{NTC_n}{\sum_{i=1}^{N_{imp}} NTC_i} \right| \tag{8}$$

where  $N_{imp}$  is the number of imperialists. To select an empire, a process similar to the roulette wheel process, which is used in selecting parents in the GA, is needed. The method of selection in the ICA is faster than the conventional roulette wheel because it is not necessary to compute the cumulative distribution function and the selection is based only on the values of possession probabilities. For this purpose, a vector of probabilities is generated as:

$$\mathbf{P} = [pp_1, pp_2, \dots, pp_{N_{imp}}] \tag{9}$$

Then the vector  $\mathbf{R}$  is created with uniformly distributed random numbers in range of [0 1]:

$$\mathbf{R} = [r_1, r_2, \dots, r_{N_{imp}}] \tag{10}$$

Finally, vector  $\mathbf{D}$  is formed as:

$$\mathbf{D} = \mathbf{P} - \mathbf{R} = [pp_1 - r_1, \dots, pp_{N_{imp}} - r_{N_{imp}}] \tag{11}$$

Referring to vector  $\mathbf{D}$  the mentioned colonies are handed to an empire whose relevant index in  $\mathbf{D}$  is maximized (Atashpaz-Gargari and Lucas 2007). Algorithm continues in several decades in this manner with the main operations (assimilation, revolution, competition) until a stop condition is satisfied. To better understand this method, the main steps of ICA are summarized in Fig. 3.

### Data input/output space

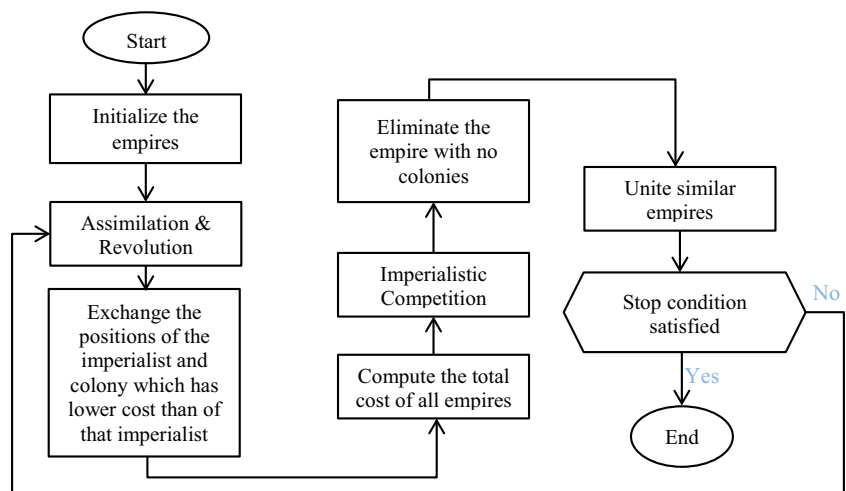
Evaluation of the constructed model is achieved based on the dataset found in the open source literature (Buckley and Morrow 2003). The developed model creates a quantitative

formulation between the input data (SARA fraction data) and the target value (crude oil RI). SARA method is a common separation technique that divides crude oil into four distinct groups namely saturate, aromatic, resin and asphaltene based on the differences in polarity and solubility. In the available dataset, fractionation is implemented through HPLC technique. In addition, crude oil RI is measured using automated index instruments GPR 11-37 refractometer. This instruments use the sodium-D line (589 nm) as a standard wavelength. The GPR 11-37 uses a critical angle technique that permits measurement of RI for opaque liquid samples. The statistical description of SARA fraction and crude oil RI data are shown in Table 1. In this study, 104 data points are used in the modeling, 30 % of which are used as validation data and the rest as training data.

### Results and discussion

In this study, an  $\epsilon$ -SVR with RBF kernel is employed for estimating the crude oil RI values from the SARA fraction data. Before training the model, all input data are normalized in range of [0 1]. The normalization task reduces confusion and increases performance of the model. The SVR parameter is optimized through three different optimization strategies. First, a common SVR optimization survey (HGP technique) is used to obtain the SVR parameters. In this strategy, a region close to the global optimal point is found with the grid search and then, the pattern search is performed over this region to find a global optimum. Second, GA and third, ICA is used to optimize the SVR model. The fitness function in the GA or its counterpart in the ICA, cost function, is introduced in the mean square error (MSE) state as follows:

Fig. 3 General flowchart of the ICA





**Table 1** Statistics of data set used in this study

Parameter	Min	Max	Mean	Standard deviation	Relationship with RI (R <sup>2</sup> )
Saturate	38.4300	78.8200	63.2527	7.6624	63.09
Aromatic	11.4800	29.7600	19.2405	3.4879	6.0e−7
Resin	0.6700	33.9200	13.9317	5.1793	62.05
Asphaltene	0.0500	16.4800	3.5730	3.3026	37.11
RI	1.4408	1.5510	1.4943	0.0185	1

**Table 2** Regulations which were used in the GA

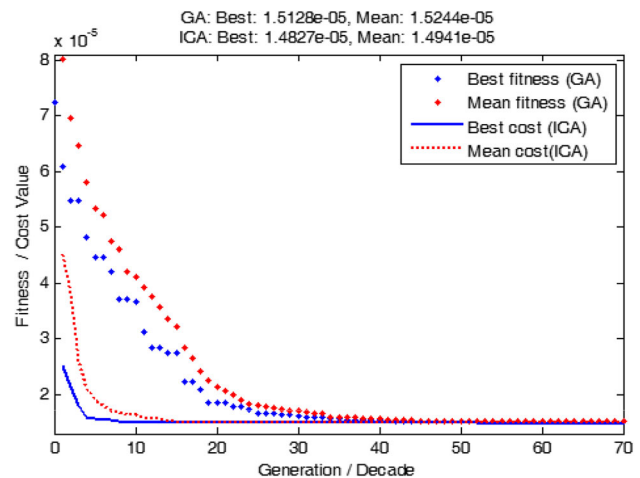
Parameter/setting	Type/value
Population type	Double vector
Population size	20 chromosomes
Initial range	[eps 1]
Scaling function	Rank
Selection function	Roulette
Elite preservation	4
Crossover function	Scatter
Crossover fraction	0.8
Mutation function	Adaptive feasible
Hybrid function	Pattern search
Max generation	70
Stall generation	70
Fitness tolerance	10 <sup>−6</sup>

**Table 3** Regulations done before run of the ICA

Parameter	Value
Number of countries	200
Number of initial imperialists	10
Number of colonies	190
Max decades	70
Range of unknown variables	[eps 50]
Revolution rate	0.3
$\xi$	0.1
$\beta$	2

$$MSE_{SVR} = \frac{1}{k} \sum_{i=1}^k (\text{output}_i - \text{target}_i)^2 \quad (12)$$

where the output is the SVR results and the target is the crude oil RI data. All regulated parameters of the GA are presented in Table 2. Also, Table 3 shows all regulations before running the ICA. The results of running the GA and the ICA for 70 iterations are shown in Fig. 4. According to this figure, the final optimum value of the ICA is 1.4941e−5 which is lower than its counterpart in the GA

**Fig. 4** Results of running ICA including best and mean costs values versus decade and Minimum and best fitness values of the GA results versus generation

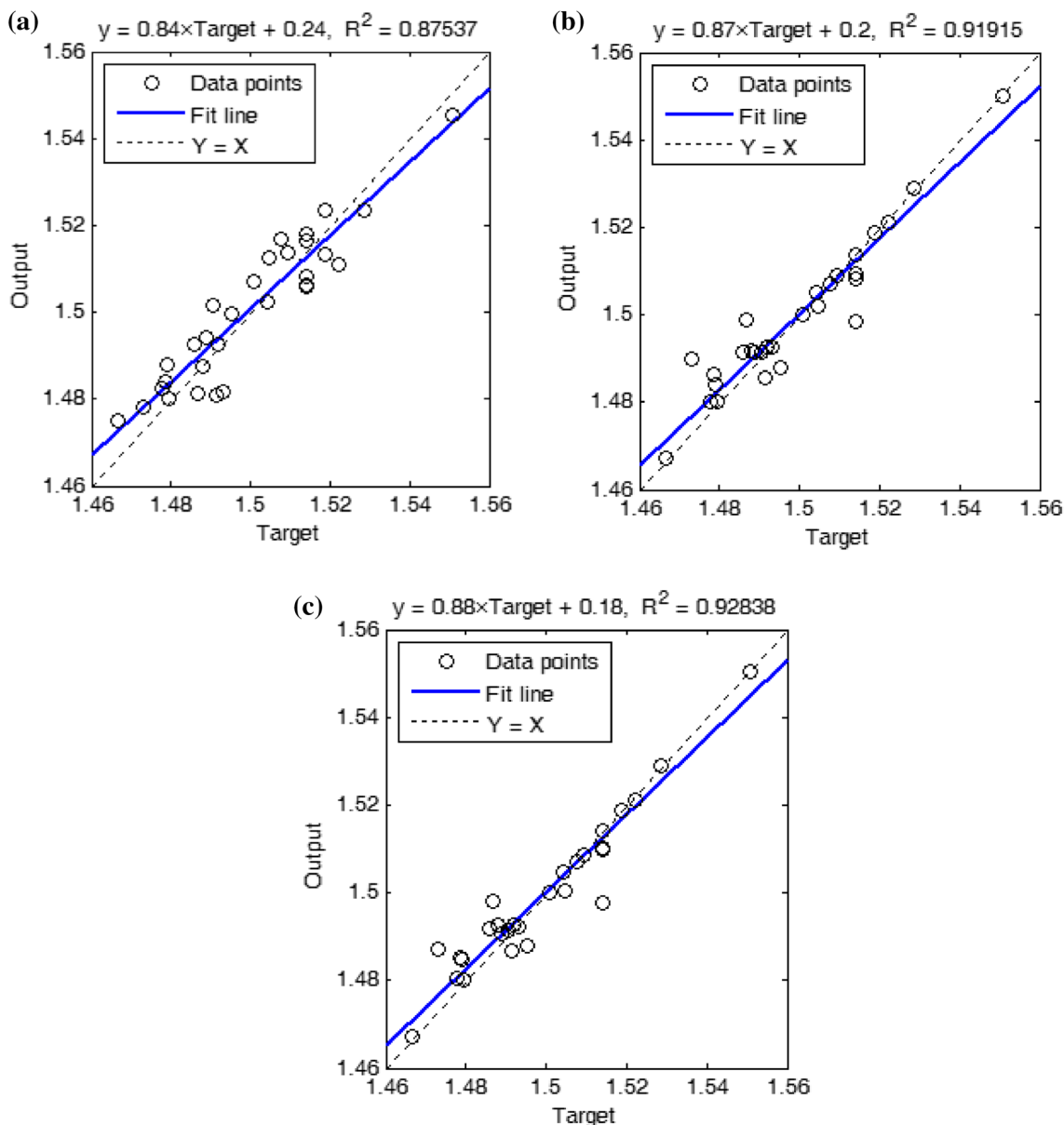
(1.5244e−5). Also in the ICA, it takes only 17 iterations for mean and best values of all imperialists to be nearly equal but in the GA, algorithm converges after 36 iterations. The corresponding values of parameters to optimize SVR model using three optimization strategies are presented in Table 4. Graphical comparison between predicted and measured crude oil RI values using HGP-SVR, GA-SVR and ICA-SVR for the validation data are shown in Fig. 5. The performance of all models is evaluated using concepts of the square of regression coefficient (R<sup>2</sup>), MSE and symmetric mean absolute percentage error (SMAPE), shown in Table 5. SMAPE is a dimensionless error measurement which expresses error as a percentage of the actual data and provides an easy way of assessing the errors. It is defined as (Flores 1986):

$$SMAPE = \frac{1}{n} \sum_{i=1}^n \frac{|\text{Predicted}_i - \text{Target}_i|}{\text{Predicted}_i + \text{Target}_i} \times 100 \quad (13)$$

According to the results presented in Table 5, the evolution algorithms including GA and ICA improve the results of SVR model. Among the presented methods, ICA-SVR decreases error up to 2.6832e−05 which corresponds to a correlation coefficient of 0.9284. Figure 6 presents

**Table 4** Optimized SVR parameters of the HGP–SVR, GA–SVR and the ICA–SVR

Model	$C$	$\varepsilon$	$\gamma$
HGP–SVR	22.725153	0.000471	0.152244
GA–SVR	7.234567	0.000490	9.981153
ICA–SVR	8.266756	0.000452	9.963511



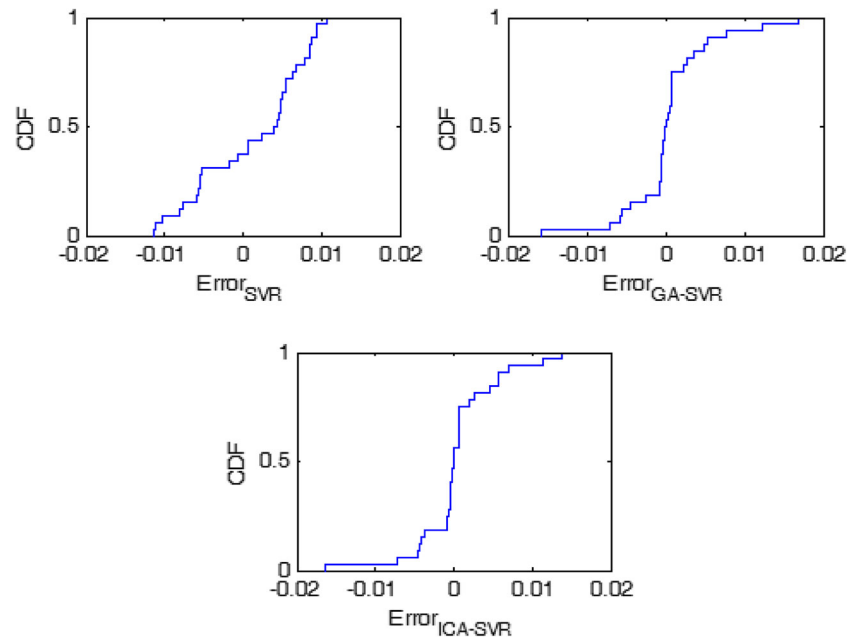
**Fig. 5** Graphical correlation between measured and predicted crude oil RI for validation data using **a** HGP–SVR, **b** GA–SVR and **c** ICA–SVR

error in terms of cumulative distribution function (CDF). The error in this figure is the difference between predicted and actual measured data. Also, details of this figure are presented in Table 6 using concept of the probability distributions of the errors. For better understanding of the

performance of the models, the error distributions of all models are presented in histograms of Fig. 7. According to this figure, 59.38 % of the error of the SVR prediction fall within one standard deviation of its mean ( $[-0.0056, 0.0081]$ ). The same analysis for GA–SVR and ICA–SVR

**Table 5** A comparisons among all methods of this study, versus  $R^2$ , MSE and SMAPE

Model	Training		Validation			
	$R^2$	MSE	SMAPE (%)	$R^2$	MSE	SMAPE (%)
HGP–SVR	0.9514	2.12156e–05	0.0833	0.8754	4.6762e–05	0.2052
GA–SVR	0.9882	1.51548e–05	0.0394	0.9192	3.0230e–05	0.1110
ICA–SVR	0.9971	1.48548e–05	0.0214	0.9284	2.6832e–05	0.1059

**Fig. 6** CDF plots of error distribution for SVR, GA–SVR and ICA–SVR**Table 6** Probability distributions of modeling errors

Model	$ \text{el}  < 0.000675$	$ \text{el}  < 0.00125$	$ \text{el}  < 0.0025$	$ \text{el}  < 0.005$	$ \text{el}  < 0.0075$
HGP–SVR	0.0938	0.0938	0.1563	0.3125	0.6250
GA–SVR	0.3438	0.5625	0.5938	0.7500	0.8750
ICA–SVR	0.4688	0.5625	0.5938	0.7813	0.9063

reveals that 78.13 % of data have an error in the range of  $[-0.0051 \ 0.006]$  and  $[-0.0049 \ 0.0056]$ , respectively. A deep insight into Figs. 6 and 7 and Table 6 suggests the superiority of the ICA–SVR when the error is decreased to zero.

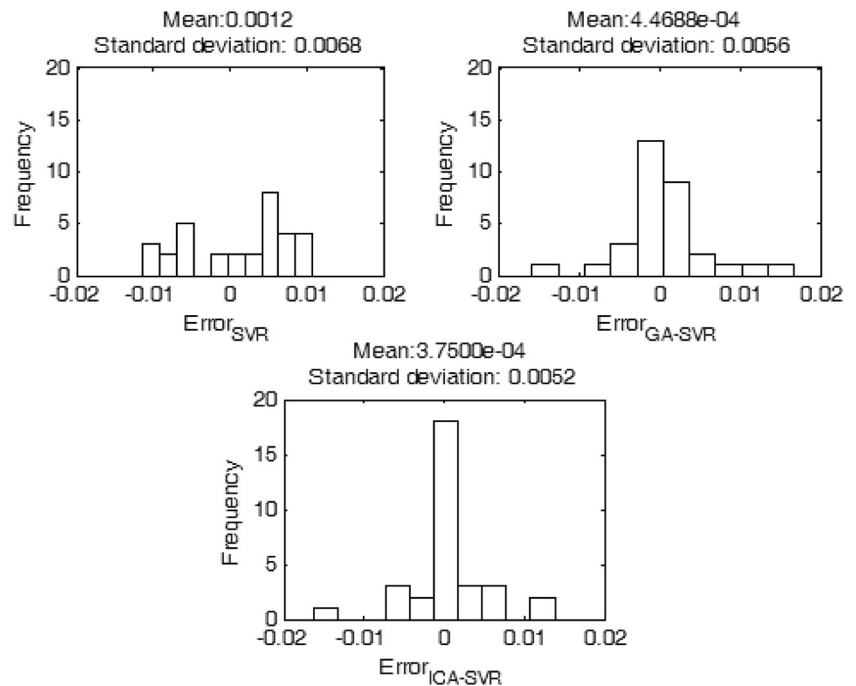
## Conclusion

Refractive index is an optical parameter of crude oil that provides valuable information about crude oil specific compositions. Hence it is crucial to develop a computer-

based approach for calculation of this property. SVR is an intelligent model used in this study for modeling of crude oil RI. Three optimization approaches namely HGP, GA, and ICA were used to find the optimal value of SVR parameters. The correlation coefficient between results of SVR optimization by HGP, GA, and ICA and the corresponding experimental values are 0.87537, 0.91915, and 0.92838, respectively. It is obvious that SVR optimized with ICA results in better prediction accuracy and outperforms the competing models (GA–SVR and HGP–SVR). This advantage is the result of ICA–SVR great performance in both rate of convergence



**Fig. 7** Error distribution of prediction from SVR, GA–SVR and ICA–SVR. A judgment based on mean and standard deviation error reveals superiority of ICA–SVR



and exploration of global optimum. The current study concluded that ICA–SVR is a promising methodology for modeling crude oil RI.

**Acknowledgments** A. Gholami would like to acknowledge the departments of research and technology of the National Iranian Oil Company and Iranian Offshore Oil Company for support throughout this research.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Afshar M, Gholami A, Asoodeh M (2014) Genetic optimization of neural network and fuzzy logic for oil bubble point pressure modeling. *Korean J Chem Eng* 31(3):496–502
- Al-Anazi AF, Gates ID (2010) Support vector regression for porosity prediction in a heterogeneous reservoir: a comparative study. *Comput Geosci* 36:1494–1503
- Al-Anazi AF, Gates ID (2012) Support vector regression to predict porosity and permeability: effect of sample size. *Comput Geosci* 39:64–76
- Alikhani-koupaei J, Abdechiri M (2010) An optimization problem for evaluation of image segmentation methods. *Int J Compu Net Security* 2:142–149
- Ansari HR (2014) Use seismic colored inversion and power law committee machine based on imperial competitive algorithm for improving porosity prediction in a heterogeneous reservoir. *J Appl Geophys* 108:61–68
- Ansari HR, Gholami A (2015a) Robust method based on optimized support vector regression for modeling of asphaltene precipitation. *J Pet Sci Eng* 135:201–205
- Ansari HR, Gholami A (2015b) An improved support vector regression model for estimation of saturation pressure of crude oils. *Fluid Phase Equilib* 402:124–132
- Ashoori S, Abedini A, Abedini R, Qorbani Nasheghi K (2010) Comparison of scaling equation with neural network model for prediction of asphaltene precipitation. *J Pet Sci Eng* 72:186–194
- Asoodeh M, Gholami A, Bagheripour P (2014a) Asphaltene precipitation of titration data modeling through committee machine with stochastically optimized fuzzy logic and optimized neural network. *Fluid Phase Equilib* 364:67–74
- Asoodeh M, Gholami A, Bagheripour P (2014b) Oil-CO<sub>2</sub> MMP determination in competition of neural network, support vector regression, and committee machine. *J Dispers Sci Technol* 35(4):564–571
- Assareh E, Behrang MA, Assari MR, Ghanbarzadeh A (2010) Application of PSO (particle swarm optimization) and GA (genetic algorithm) techniques on demand estimation of oil in Iran. *Energy* 35(12):5223–5229
- Atashpaz-Gargari E, Lucas C (2007) Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition. In: *IEEE paper*
- Atashpaz-Gargari E, Hashemzadeh F, Rajabioun R, Lucas C (2008) Colonial competitive algorithm: a novel approach for PID controller design in MIMO distillation column process. *Int J Intel Comput Cybern* 1:337–355
- Bagheripour P, Gholami A, Asoodeh M, Vaezzadeh-Asadi M (2015) Support vector regression based determination of shear wave velocity. *J Pet Sci Eng* 125:95–99
- Buckley JS, Morrow NR (2003) Wettability and imbibition: microscopic distribution of wetting and its consequences at the core and field scales. Final Report. submitted by: New Mexico Petroleum Recovery Research Center
- Buckley JS, Wang J (2002) Crude oil and asphaltene characterization for prediction of wetting alteration. *J Pet Sci Eng* 33:195–202

- Chamkalani A (2012) Correlations between SARA fractions, density and RI to investigate the stability of asphaltene. *ISRN Anal Chem*. doi:10.5402/2012/219276
- Chamkalani A, Mohammadi AH, Eslamimanesh A, Gharagheizi F, Richon D (2012) Diagnosis of asphaltene stability in crude oil through “two parameters” SVM model. *Chem Eng Sci* 81:202–208
- Chamkalani A, Chamkalani R, Mohammadi AH (2014) hybrid of two heuristic optimizations with LSSVM to predict refractive index as asphaltene stability identifier. *J Dispers Sci Technol* 35:1041–1050
- El Ghandoor H, Hegazi E, Nasser I, Behery GM (2003) Measuring the refractive index of crude oil using a capillary tube interferometer. *Opt Laser Technol* 35:361–367
- Fan T, Wang J, Buckley JS (2002) Evaluating crude oils by SARA analysis. *SPE paper* 75228
- Flores BE (1986) A pragmatic view of accuracy measurement in forecasting. *Omega* 14:93–98
- Gholami A, Asoodeh M, Bagheripour P (2014a) Fuzzy assessment of asphaltene stability in crude oils. *J Dispers Sci Technol* 35:1041–1050
- Gholami A, Asoodeh M, Bagheripour P (2014b) Smart Determination of difference index for asphaltene stability evaluation. *J Dispers Sci Technol* 35:572–576
- Gholami A, Asoodeh M, Bagheripour P (2014c) How committee machine with SVR and ACE estimates bubble point pressure of crudes. *Fluid Phase Equilib* 382:139–149
- Gholami A, Afshar M, Bagheripour P, Asoodeh M, Vaezzadeh-Asadi M (2015) Smart correlation of compositional data to saturation pressure. *J Nat Gas Sci Eng* 22:661–669
- Hemmati-Sarapardeh A, Khishvand M, Naseri A, Mohammadi AH (2013) Toward reservoir oil viscosity correlation. *Chem Eng Sci* 90:53–68
- Holland JH (1975) *Adaptation in natural and artificial systems*. University of Michigan Press, Michigan
- Khabbazi A, Atashpaz-Gargari E, Lucas C (2009) Imperialist competitive algorithm for minimum bit error rate beam forming. *Int J Bio-Inspired Comput* 1:125–133
- Khishvand M, Khamehchi E (2012) Nonlinear risk optimization approach to gas lift allocation optimization. *Ind Eng Chem Res* 51(6):2637–2643
- Mohaghegh S (2000) Virtual-intelligence applications in petroleum engineering: part 2-Evolutionary Computing. *J Pet Technol* 52:40–46
- Naseri A, Khishvand M, Sheikhloo AA (2014) A correlations approach for prediction of PVT properties of reservoir oils. *Pet Sci Technol* 32(17):2123–2136
- Rajabioun R, Atashpaz-Gargari E, Lucas C (2008) Colonial competitive algorithm as a tool for Nash equilibrium point achievement. *Comput Sci Appl* 5073:680–695
- Rasuli Nokandeh N, Khishvand M, Naseri A (2012) An artificial neural network approach to predict asphaltene deposition test result. *Fluid Phase Equilib* 329:32–41
- Riazi MR, Al-Otaibi GN (2001) Estimation of viscosity of liquid hydrocarbon systems. *Fuel* 80:27–32
- Shokooh Saljooghi B, Hezarkhani A (2014) Comparison of WAVENET and ANN for predicting the porosity obtained from well log data. *J Pet Sci Eng* 123:172–182
- Shokooh Saljooghi B, Hezarkhani A (2015) A new approach to improve permeability prediction of petroleum reservoirs using neural network adaptive wavelet (wavenet). *J Pet Sci Eng* 133:851–861
- Steinwart I, Christmann A (2008) *Support vector machines*. Springer, New York
- Tatar A, Shokrollahi A, Halali MA, Azari V, Safari H (2015) A Hybrid intelligent computational scheme for determination of refractive index of crude oil using SARA fraction analysis. *Can J Chem Eng* 93(9):1547–1555
- Taylor SD, Czarnecki J, Masliyah J (2001) Refractive index measurements of diluted bitumen solutions. *Fuel* 80:2013–2018
- Touba H, Mansoori GA, Sarem AMS (1997) New analytic techniques for petroleum fluid characterization using molar refraction. *SPE paper* 38312
- Vapnik V (1982) *Estimation of dependences based on empirical data*. Springer, Berlin
- Vapnik V (1995) *The nature of statistical learning theory*. Springer, New York
- Vapnik V (1998) *Statistical learning theory*. Wiley, New York
- Vargas FM, Chapman WG (2010) Application of the one-third rule in hydrocarbon and crude oil systems. *Fluid Phase Equilib* 290:103–108
- Vert JP, Schölkopf B, Tsuda K (2004) *Kernel methods in computational biology*. MIT Press, Cambridge
- Zargar G, Gholami A, Asoodeh M, Bagheripour P, Vaezzadeh-Asadi M (2015) PSO-fuzzy eliminates deficiency of neuro-fuzzy in assessment of asphaltene stability. *Indian J Chem Technol* 22:135–140