ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

Refractive Index of Hemoglobin Analysis: A Comparison of Alternating Conditional Expectations and Computational Intelligence Models

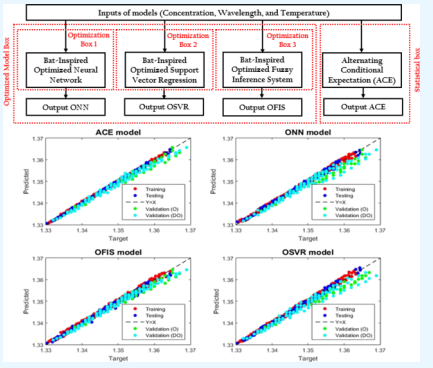
Aida Alizamir, Amin Gholami, Nader Bahrami, and Mehdi Ostadhassan*

Cite This: *ACS Omega* 2022, 7, 33769–33782

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Hemoglobin is one of the most important blood elements, and its optical properties will determine all other optical properties of human blood. Since the refractive index (RI) of hemoglobin plays a vital role as a non-invasive indicator of some illnesses, accurate calculation of it would be of great importance. Moreover, measurement of the RI of hemoglobin in the laboratory is time-consuming and expensive; thus, developing a smart approach to estimate this parameter is necessary. In this research, four viable strategies were used to make a quantitative correlation between the RI of hemoglobin and its influencing parameters including the concentration, wavelength, and temperature. First, alternating conditional expectations (ACE), a statistical approach, was employed to generate a correlation to predict the RI of hemoglobin. Then, three different optimized intelligent techniques—optimized neural network (ONN), optimized fuzzy inference system (OFIS), and optimized support vector regression (OSVR)—were used to model the RI. A bat-inspired (BA) algorithm was embedded in the formulation of intelligent models to obtain the optimal values of weights and biases of an artificial neural network, membership functions of the fuzzy inference system, and free parameters of support vector regression. The coefficient of determination, root-mean-square error, average absolute relative error, and symmetric mean absolute percentage error for each of the ACE, ONN, OFIS, and OSVR were found as the measure of each model's accuracy. Results showed that ACE and optimized models (ONN, OFIS, and OSVR) have promising results in the estimation of hemoglobin's RI. Collectively, ACE outperformed ONN, OFIS, and OSVR, while sensitivity analysis indicated that the concentration, wavelength, and, lastly, temperature would have the highest impact on the RI.



1. INTRODUCTION

Hemoglobin is a vital protein in human blood cells that is responsible for carrying oxygen between the lungs and the rest of the tissues in the human body. This process will allow aerobic respiration to provide energy to the cells for their metabolism.[1,2] There are two distinct forms of hemoglobin, that is, hemoglobin exist in oxygenated (O) and deoxygenated (DO) states based on their capability to reversibly bind up to four oxygen molecules.[1] Besides, hemoglobin is a biological marker to clinically diagnose different medical conditions mainly because of the close relationship between the pathophysiology of various diseases and the red blood cells.[3−6] Considering the shape and size of hemoglobin, its optical properties would significantly affect the optical attributes of the entire blood.[7] Thus, optical properties of hemoglobin can provide us with valuable information which is extremely useful in medical diagnosis[4] and the patient's overall health conditions. At the same time, its role in therapeutic applications, particularly laser medicine, is highly valuable. Moreover, the RI of hemoglobin, which can be found through optical methods in the laboratory using a refractometer, can reveal the true nature of the blood cell disorder.[8−10] However, experimental measurement of this parameter in the laboratory is highly expensive, labor-intensive, and time consuming. This has promoted the development of empirical models based on statistical approaches instead to obtain the RI more rapidly,[11] while many of such models have demonstrated a poor performance.

Based on what was said above, machine learning (ML), which has been proven to be an efficient method to replace common statistical approaches and a new modeling tool in solving complicated regression and classification problems, could be used to estimate the RI of hemoglobin as well. In this realm, in the classification task, a number of studies were conducted to successfully classify the blood cells.[12−17] Tomari et al.(2014) used an artificial neural network (ANN) to classify red blood cells as normal/abnormal.[12] Kultu et al. (2020) proposed a convolutional neural network (CNN) to identify and locate white blood cell types in blood images, which led to an increase in the performance of existing blood test devices.[13] Toğaçar et al.

Received: February 5, 2022
Accepted: August 30, 2022
Published: September 13, 2022

(2020) adopted CNN models to improve the classification success of white blood cell types.[14] To address the issue of multiple cell overlap in white blood cell image classification, Patil et al. (2021) combined the deep learning approach (merging of the CNN and recurrent neural network model) with canonical correlation analysis.[15] Girdhar et al. (2022) also employed the CNN model to classify white blood cell.[16] Davamani et al. (2022) developed fuzzy c-means clustering for blood cell classification.[17] In addition, others have developed AI-based models to segment blood vessels.[18−25] Wang et al. (2015) integrated CNN and random forest (RF) for retinal blood vessel segmentation.[19] Soomro et al. (2019) utilized deep CNN for segmenting retinal blood vessels.[20] In their paper, in order to generate contrast images for training data sets, morphological mappings along with the principal component analysis-based pre-processing steps were used. Zhang et al. (2020) used a U-net-based deep learning approach to track and segment brain blood vessels in digital subtraction angiography images.[21] Tchinda et al. (2021) proposed a potent strategy based on classical edge detection filters and ANNs for the segmentation of blood vessels in retinal photographs.[22] In their paper, edge detection filters were first applied to extract the feature vector. Then, the resulting features were used to train an ANN in order to recognize each pixel as belonging to blood vessels or not. Gegundez-Arias et al. (2021) developed a new deep learning method for blood vessel segmentation in retinal images based on convolutional kernels and a modified U-net model.[23] Deng and Ye (2022) performed retinal blood vessel segmentation based on an improved deformable convolutional M-shaped network and a pulse-coupled neural network.[24] Zhang et al. (2022) presented a novel automatic method based on bridge-net by joint learning context-involved and non-context features for the segmentation of retinal blood vessels.[25] In the regression task, researchers have attempted to develop intelligence-based approaches to reliably predict blood pressure.[26−31] Xu et al. (2017) presented a capable methodology based on ANN for continuous blood pressure estimation based on multiple parameters from electrocardiogram and photoplethysmogram.[26] Senturk et al. (2020) constructed a predictive model based on dynamic recurrent neural networks to estimate non-invasive continuous cuffless blood pressure.[27] Esmaelpoor et al. (2020) proposed a two-step strategy for blood pressure estimation using photoplethysmogram signals.[28] In the first stage, they employed CNN to extract morphological features from each photoplethysmogram segment and then to estimate systolic and diastolic blood pressure separately. In the second stage, they used long short-term memory (LSTM) to capture temporal dependencies. Baker et al. (2021) applied a hybrid neural network for continuous and non-invasive estimation of blood pressure from raw electrocardiogram and photoplethysmogram waveforms.[29] Qiu et al. (2021) proposed a hybrid neural network architecture, which contained a CNN-Sequential-Adapt layer, a ResNet25_BP layer with squeeze and excitation blocks and fully connected layers, for blood pressure estimation.[30] Cheng et al. (2021) employed fully CNN for prediction of arterial blood pressure waveforms from photoplethysmogram signals.[31] Others attempted to carefully evaluate the applicability of ML in the estimation of blood glucose levels.[32−36] Ben Ali et al. (2018) used ANN for continuous blood glucose level prediction of type 1 diabetes.[32] D'Antoni et al. (2020) constructed an auto-regressive time-delayed jump neural network for blood glucose level prediction.[33] Alfian et al. (2020) applied ANN, support vector regression, K-nearest neighbor, decision tree, RF, adaptive boosting, and extreme gradient boosting models for blood glucose prediction of type 1 diabetes.[34] They compared results of predictive models and concluded that ANN outperforms other intelligence-based models. Dudukcu et al. (2021) predicted blood glucose by virtue of deep neural networks. LSTM, WaveNet, and gated recurrent units, and decision-level combinations of these architectures were deep learning methods which were used to predict blood glucose.[35] Zhang et al. (2021) adopted deep learning and regression approaches to forecast blood glucose levels for type 1 diabetes.[36] Four data-driven models including different neural network architectures, a reservoir computing model, and a linear regression approach were employed in their study. Considering the RI estimation of hemoglobin, AI methods have shown some promising results compared to common statistical approaches.[37,38] All in all, these studies confirm the applicability of ML and AI methods in the analysis of blood from various perspectives, while the estimation of RI, despite its importance, has been rarely done via intelligent techniques.

In this study, four techniques are presented to estimate the RI of hemoglobin. To do so, alternating conditional expectation (ACE) was first employed to carry out a quantitative estimation of the RI of hemoglobin. This method correlates inputs to outputs through optimally transferring dependent and independent variables into high-dimensional space. Next, three optimized models, viz., the optimized neural network (ONN), optimized fuzzy inference system (OFIS), and optimized support vector regression (OSVR) were used to model the RI. In neural network equations, a bat-inspired algorithm (BA) was embedded to achieve global minima (optimal value of weight and bias), where the neural network can operate at its maximum capacity. In the fuzzy inference system, BA was employed to optimize membership functions of a fuzzy inference system established in the structure of the Sugeno fuzzy inference system type. Moreover, BA was involved in the support vector regression equations to improve its precision by means of determining the optimal values of free parameters. Finally, sensitivity analysis (SA) was performed to obtain information related to the contribution of each input in every model. This SA will enable us to understand how each parameter would affect the outputs of every model.

2. MODEL DESCRIPTION

2.1. Alternating Conditional Expectations. Recently, Breiman and Freidman formulated a nonparametric approach known as ACE for solving complex regression problems.[39] When the functional form between dependent variables and independent ones is implicit, this method is far superior to its rivals in identifying principal nonlinear relationships of input/output space.[40−43] This exceptional feature motivated scientists to use ACE in order to find solutions to complicated problems which were impossible to be solved by the virtue of conventional regression methods.[39] A full description of ACE can be found in the original paper of Breiman and Freidman,[39] but in general a linear regression model for p independent variables X_1, X_2,....., X_p and a response variable Y can be mathematically shown as follows.[39]

$$Y = \delta_0 + \sum_{i=1}^{p} \delta_i X_i + \varepsilon \quad (1)$$

https://doi.org/10.1021/acsomega.2c00746
ACS Omega 2022, 7, 33769−33782

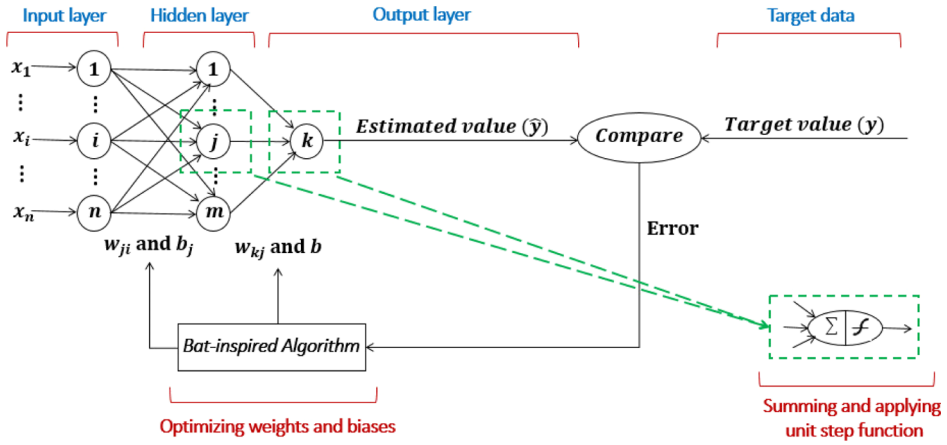


Figure 1. Schematic diagram of the BA optimally determining the weights and the bias of the neural network.

In the above equation, regression coefficients ($\delta_i, i = 0-p$) and the error term (ε) are determined during the regression analysis. Compared to the linear regression method, ACE transforms dependent and independent variables into a high-dimensional space and solves the following regression equation in those spaces.[39]

$$\vartheta(Y) = \alpha + \sum_{i=1}^{p} \xi(X_i) + \varepsilon \tag{2}$$

where $\vartheta(Y)$ and $\xi_1(X_1),....,\xi_p(X_p)$ are arbitrary measurable mean-zero functions of the Y, X_1, X_2,...., X_p, respectively. For achieving the best model via ACE, optimal transformation (OT) $\xi_1^*(X_i), i = 1,...,p$ and $\vartheta^*(Y)$ which changes the dimension of the data input/output space must be extracted. In this desired space, there is a maximum correlation between transformed dependent variables and the sum of the transformed predicted variables. With the above objective in consideration, the following equation must be minimized.[39]

$$\varepsilon^2(\vartheta, \xi_i, \ldots., \xi_p) = E\left(\left[\vartheta(Y) - \sum_{i=1}^{p} \xi(X_i)\right]\right)^2 / E\vartheta^2(Y) \tag{3}$$

In order to minimize the above equation, Breiman and Freidman took advantage of a series of single-function minimizations and proposed the following equations for the response variable and predictor, respectively:[39]

$$\vartheta(Y) = E\left[\sum_{i=1}^{p} \xi(X_i) \middle| Y\right] \Big/ \left\| E\left[\sum_{i=1}^{p} \xi(X_i) \middle| Y\right] \right\| \tag{4}$$

$$\xi_{j,1}(X_j) = E\left[\vartheta(Y) - \sum_{i=1}^{p} \xi_i(X_i) | X_k\right] \tag{5}$$

By iteratively changing the values of the eqs 4 and 5, optimal transformation $\xi_1^*(X_i), i = 1,...,p$, and $\vartheta^*(Y)$, where the value of ε^2 in eq 3 has the minimum values, will be extracted. In the desired space, the response value is related to the predictor variables with the following equation.[39]

$$\vartheta^*(Y) = \sum_{i=1}^{p} \xi_i^*(X_i) + e^* \tag{6}$$

Here, e* is an undesirable error that ACE is unable to capture.[39]

2.2. Optimized Neural Network. A neural network is an algorithm that is based on biological neural networks of the human brain, and it has been proven useful for modeling intricate regression tasks.[44] Exactly analogous to the human brain, neural networks are incredibly intelligent to effectively train from a number of observed data in the absence of an initially defined mathematical relationship between dependent and independent variables. Many scientists and engineers have recently been applying this mathematical technique to find out the complicated relationships of the input/output space.[45−49] Considering its capability, it is a prospective candidate to find solutions for nonlinear problems.[50−53] Three layers are defined in the structure of a neural network, including input, hidden, and output layers. The input layer is devoted to receiving the input data, while output layers are used to generate the final result. The hidden layer is employed to obtain the dependence of output data on the input data. For the training task, this approach calculates the value of output from the input data through the entire network and then determines the difference between estimated values and the corresponding observed values. Subsequently, this difference is propagated backward through the network as an error, thereby adjusting weights and biases. The training phase is considered fulfilled when the error reaches a minimum, which yields the best predictive model. Although the neural network is a promising and applicable technique to model nonlinear problems, it may pose some challenges in distinguishing the global minima from the local minima and can produce misleading results.[54−57] To effectively solve this issue, scientists have improved the performance of neural networks through hybridization of optimization algorithms and this intelligent model.[58−60] Considering this solution, in this research, the structure of neural networks incorporates BA to find out optimal values of their weights and biases where the problem is preferably trapped in the global minima. A schematic diagram of the BA optimally determining the weights and the bias of a neural network is illustrated in Figure 1. In this study, ONN was generated via MATLAB programming.

2.3. Optimized Fuzzy Inference System. A fuzzy inference system is a methodology established from Zadeh's fuzzy set theory, which has been proved notably practical in modeling nonlinear systems.[61] This approach enables scientists with mathematical computation to handle uncertainty in accordance with "degrees of truth" in lieu of "true or false," which is generally carried out in the logic of the calculation. In this approach, each membership function represents a fuzzy set.

https://doi.org/10.1021/acsomega.2c00746
ACS Omega 2022, 7, 33769−33782

A fuzzifier, an inference engine (or a fuzzy rule base), and the defuzzifier are three principal constituents of fuzzy inference systems. The fuzzifier employs a membership function to map a set of data inputs to intervals of 0 and 1. The purpose of the inference engine as a second part of the fuzzy inference system is to apply fuzzy rules to the converted output data of the fuzzifier. Outputs of the inference engine are integrated into one fuzzy output distribution and eventually transferred into a crisp output by virtue of the output membership function in the defuzzifier process.[62] In the current work, the Gaussian membership function is used as an input, and a linear polynomial function is employed as an output membership function. The two most capable types of fuzzy inference systems are Mamdani and Sugeno, where Sugeno is utilized to perform the modeling task in this study. The Sugeno model is composed of "if−then" rules in the following general form[62]

$$R_n\text{: } I_1 \text{ is } M_n^1 \text{ and } I_2 \text{ is } M_n^2 \text{ and } \ldots \ I_{\mathrm{m}} \text{ is } M_n^m \text{ then } Z_n = f_n(I) \quad (7)$$

where n refers to the number of the rules; I_i is the input of the fuzzy inference system; $M_{\mathrm{n}}^{\mathrm{m}}$ is the Gaussian input membership function of mth input data and nth rule, and $Z_n = f_n$ (I) is a function in the series. As mentioned above, the output membership function f_n (I) is a linear polynomial in the input variables. The output level, Z_n, is weighted by the firing strength for each rule, and finally, the overall output is estimated via a weighted average operator. One of the most important tasks in modeling a regression problem is to find the optimum values of membership functions since these values play a critical role in the accuracy of the generated model.[63,64] Therefore, to increase the efficiency of the fuzzy inference system in mapping the functional dependency between RI and its input variables, both BA and the fuzzy inference system were integrated in this study to reach the best values for the membership function. The main idea of implementing such a combination is presented in Gholami et al.[56] In the current paper, OFIS was built in the MATLAB environment.

2.4. Optimized Support Vector Regression. Support vector regression, which was proposed by Vapnik, is a type of intelligent model based on statistical theory to drive equations for relating input variables and outputs.[65] The estimation approach in this function is a refinement of a support vector machine to solve knotty regression tasks. Recently, this method has attracted much attention because it showed significant ability in solving diverse regression problems with excellent accuracy.[66−69] In comparison to neural networks, this technique is superior in generalization domains which originated from how it minimizes the risk. In this regard, support vector regression applies the structural risk minimization principle to reduce the upper bound on expected risk, while neural networks use the empirical risk minimization principle to decrease the error on the training data.[70] In order to transfer nonlinear learning problems into linear ones, support vector regression utilizes kernel function in its formulation. Based on the literature, four dominant kernel functions that are generally used for this transformation are linear, polynomial, radial basis function, and sigmoid. However, it has been found that the radial basis function is a better kernel to be used in regression problems. Since the influence of penalty parameters of support vector regression on the performance of a constructed model is greatly high, it is considered as one of the most challenging issues while training the model.[71−83] Thus, finding the values of these parameters plays a key role in successfully applying this method for solving knotty problems. In this study, BA was combined with support vector regression to improve its efficiency by calculating the optimum value of free parameters while OSVR was constructed in the MATLAB environment.

2.5. Bat-Inspired Algorithm. Yang developed the BA as the most modern meta-heuristic algorithm with the ability to perform a global search to manage different sets of optimization problems.[84] The purpose of this method is to find the best solution to the optimization problem by simulation of bats' behavior in finding food and prey. Bats use their hearing ability to estimate the location of their surrounding matters, which is the foundation of establishing the BA method.[84] Many scientists conducted research to compare the simplicity, speed, and ability of BA with other optimization tasks.[74,78] Results show that BA performs better than other approaches. This is due to the method's excellent potential to substantially decrease the error encountered with the estimation.[78] In this paper, BA was used as an optimization algorithm to enhance the efficiency of the intelligent models by finding the optimal values of those parameters.

3. DATA INPUT/OUTPUT SPACE

To build and evaluate the predictive models in this paper, experimental data of hemoglobin's RI of humans are obtained from Yahya and Saghir.[11] They explain that since it is difficult to keep real human blood samples intractable (a condition when the samples have homogeneous hemoglobin concentration and optical stability), other samples resembling the blood from freeze-dried human hemoglobin powder dissolved in phosphate buffered saline (PBS) were utilized as well. To get blood solutions, a specific mass of lyophilized blood was dissolved in PBS which acted as a solvent. PBS keeps the pH of the blood solutions at 7.4, which is very important since the RI is affected by changes in the pH. A precise Peltier thermostat and a multiwavelength Anton Paar WR refractometer were used to measure the temperature variations and the RI of human blood samples of different wavelengths (436.1−657.2 nm), respectively. Snell's law is the basic principle of the device, and it specifies the critical angle of the total internal reflection of the investigated sample. A sensor array was employed to detect the projected beam, so the RI of blood samples can be calculated. In fact, three selected features of concentration, wavelength, and temperature were considered as input data to construct the correlation.[11] It is important to mention that 80% (384 sample points) of the data points were selected to train the model and the remaining 20% (96 sample points) as the test data set to examine the performance of the constructed model. In addition to this data set, two other data sets from Zhernovaya et al.[85] were utilized to check the validation of the developed models. Each data set has 126 data points. In Zhernovaya et al.,[85] Sigma-Aldrich's human hemoglobin (lyophilized powder) was used to obtain experimental results. Similar to Yahya and Saghir,[11] PBS is employed to dissolve human hemoglobin to keep it at a pH of 7.4 and to avoid changes in the RI from pH variations. To measure RI, the digital multiwavelength refractometer DSR-λ was employed where the RI was obtained by the measurements of the angle of total internal reflection.[11] Measurement was conducted for solutions of DO and O hemoglobin, which were obtained from adding sodium dithionite and sodium bicarbonate, respectively, to all samples. Statistical analysis data for input and output parameters, including minimum, maximum, and mean value, mode, standard deviation, skewness, kurtosis,

https://doi.org/10.1021/acsomega.2c00746
ACS Omega 2022, 7, 33769−33782

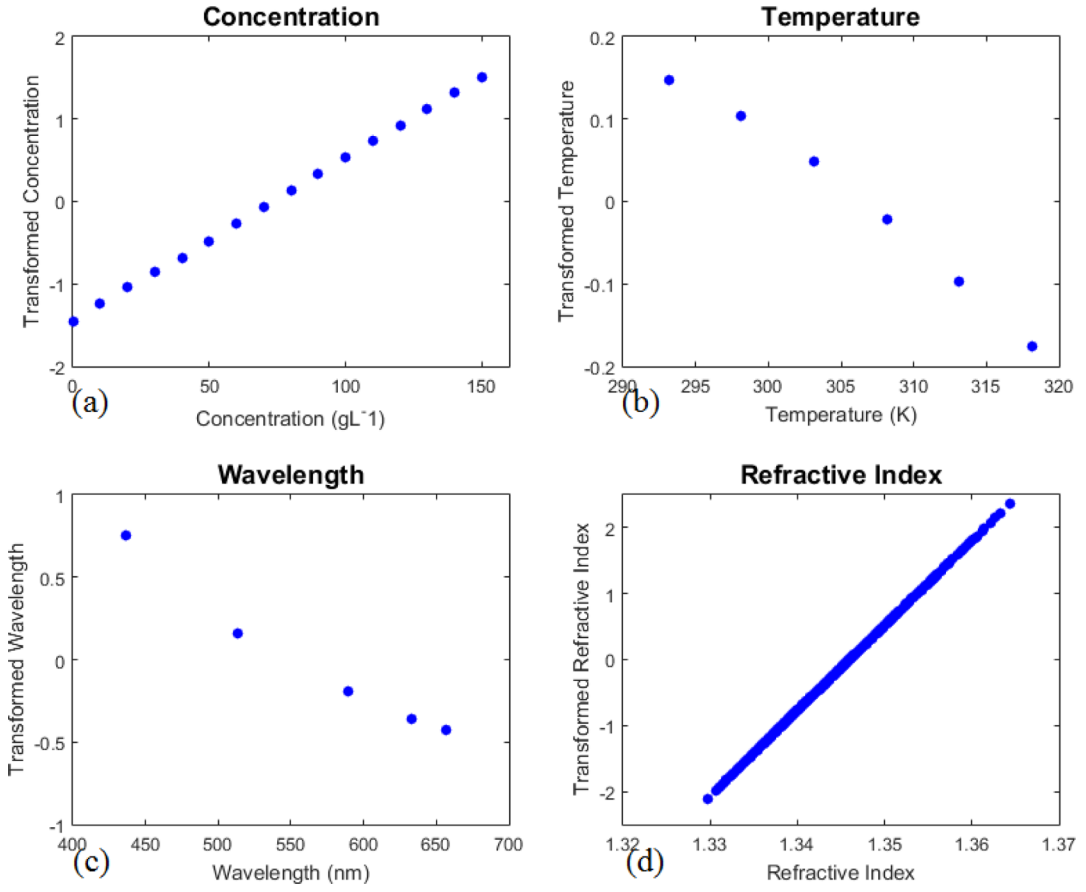


Figure 2. Optimal transformation of (a) concentration, (b) temperature, (c) wavelength, and (d) RI. Via curve fitting, equations that relate the value of optimal transformation of each parameter with its value were mathematically obtained.

and coefficient of variation of collected data, is tabulated in Table S1 of Supporting Information.

4. RESULTS AND DISCUSSION

In this study, the coefficient of determination (R^2), mean square error (MSE), average absolute relative error (AARE), and symmetric mean absolute percentage error (SMAPE) were used to make decisions about the accuracy of the developed models. Their expressions are given in eqs 8−11 as follows:

Table 1. Polynomial Coefficients for Determining Optimal Transformation of Each Input Variable

L	B_0	B_1	B_2
concentration	−1.4511274748	0.0197881143	0
wavelength	7.1323992743	−0.0208643386	0.0000142602
temperature	4.0046826041	−0.0130999526	0

Coefficient of determination (R^2)[54]

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (X_{\text{i pred}} - X_{\text{i obs}})^2}{\sum_{i=1}^{n} (X_{\text{i pred}} - \overline{X}_{\text{obs}})^2} \quad (8)$$

MSE[54]

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (X_{\text{i obs}} - X_{\text{i pred}})^2 \quad (9)$$

AARE[54]

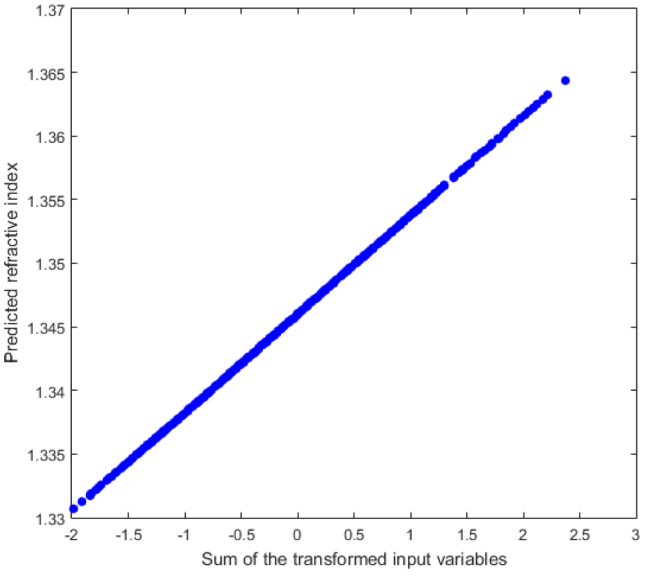


Figure 3. Relationship between the estimated RI and the sum of the transformed input variables. This linear function leads to the calculation of hemoglobin's RI from optimal transformation of independent variables.

$$\text{AARE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{X_{\text{i obs}} - X_{\text{i pred}}}{X_{\text{i obs}}} \right| \quad (10)$$

SMAPE[79]

https://doi.org/10.1021/acsomega.2c00746
ACS Omega 2022, 7, 33769−33782

Table 2. Polynomial Coefficients of the Archived ACE Model for Estimation of RI

parameter	value
α	0.000153666146350
β	−0.000101861894337
δ	−0.000162023650231
λ	0.000000110738826
μ	1.421225072812270

$$\text{SMAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{X_{\text{i obs}} - X_{\text{i pred}}}{X_{\text{i obs}} + X_{\text{i pred}}}\right| \quad (11)$$

where $Y_{\text{i obs}}$ is the measured value of sample i, $Y_{\text{i pred}}$ is the predicted value of sample i, $\hat{Y}_{\text{obs}}$ is the average of the measured value, and n is the number of samples. When the values of AARE, SMAPE, and MSE are close to 0, as well as the value of R^2 is close to 1, a model with superb performance is achieved. It is worth mentioning that in the training stage of the constructed optimized models (ONN, OFIS, and OSVR), a fourfold cross-validation technique was employed to train intelligence-based models and generate predictive models which produce stable results.

4.1. ACE Results. The ACE method was used to generate a model to quantitatively estimate the RI of hemoglobin. To achieve this, ACE optimally transfers input parameters and target values to a high-dimensional space. Figure 2 illustrates the optimal transformations for input and output variables of training data points (384 data points). In this high-dimensional space, transformed dependent variables and the sum of transformed predicted variables have a maximum correlation. Based on Figure 2, we can understand how the optimal transformation of each variable is related to its value.

To construct ACE models, equations relating each independent parameter to its transformation must be carefully extracted. In this study, this task was performed through a curve-fitting toolbox in the MATLAB environment. In this study, developed equations which elicit identical transformed values of each parameter are simple polynomial expressions. These mathematical equations for independent parameters (concentration (C), wavelength (W), and temperature (T)) are given as follows

$$\text{Tr}(L) = \sum_{i=0}^{2} B_i L^i \quad (12)$$

where L is the value of non-transformed input of a model and Tr is the optimal transformation of the L parameter. Coefficients that were determined through a curve-fitting tool and were extracted for calculating optimal transformations of each independent variable are tabulated in Table 1.

In fact, eq 12 maps inputs to a desired high-dimensional space where transformed dependent variables and the sum of the transformed predicted variables have a maximum correlation. The following equation provides the summation value of input's optimal transformations.

$$\text{S(Tr)} = \text{Tr(C)} + \text{Tr(W)} + \text{Tr(T)} \quad (13)$$

In which, Tr ©, Tr (W), Tr (T), and S (Tr) are optimal transformation of concentration, optimal transformation of

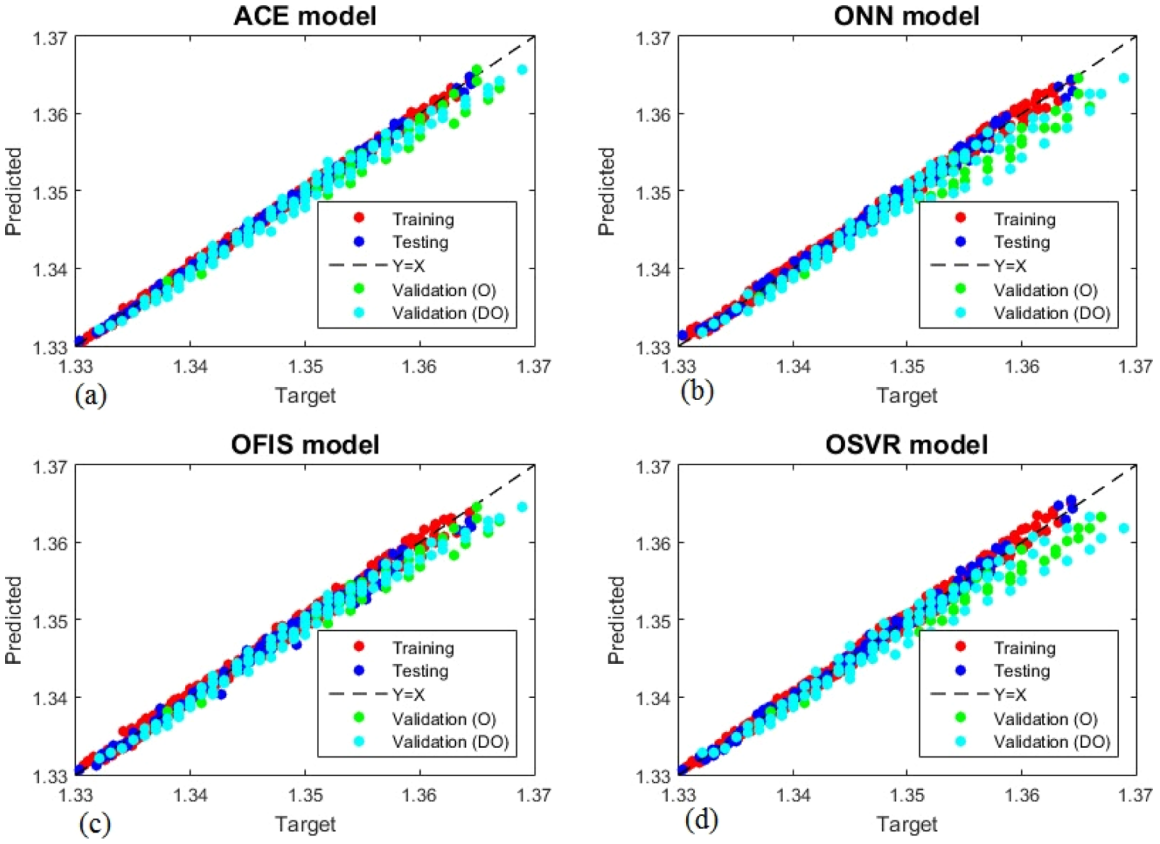


Figure 4. Cross-plot of estimated vs measured values for (a) ACE, (b) ONN, (c) OFIS, and (d) OSVR. This figure was plotted for training, testing, and two validation data sets. Based on this figure, it can be concluded that constructed models have acceptable fitness with experimental data.

https://doi.org/10.1021/acsomega.2c00746
ACS Omega 2022, 7, 33769−33782

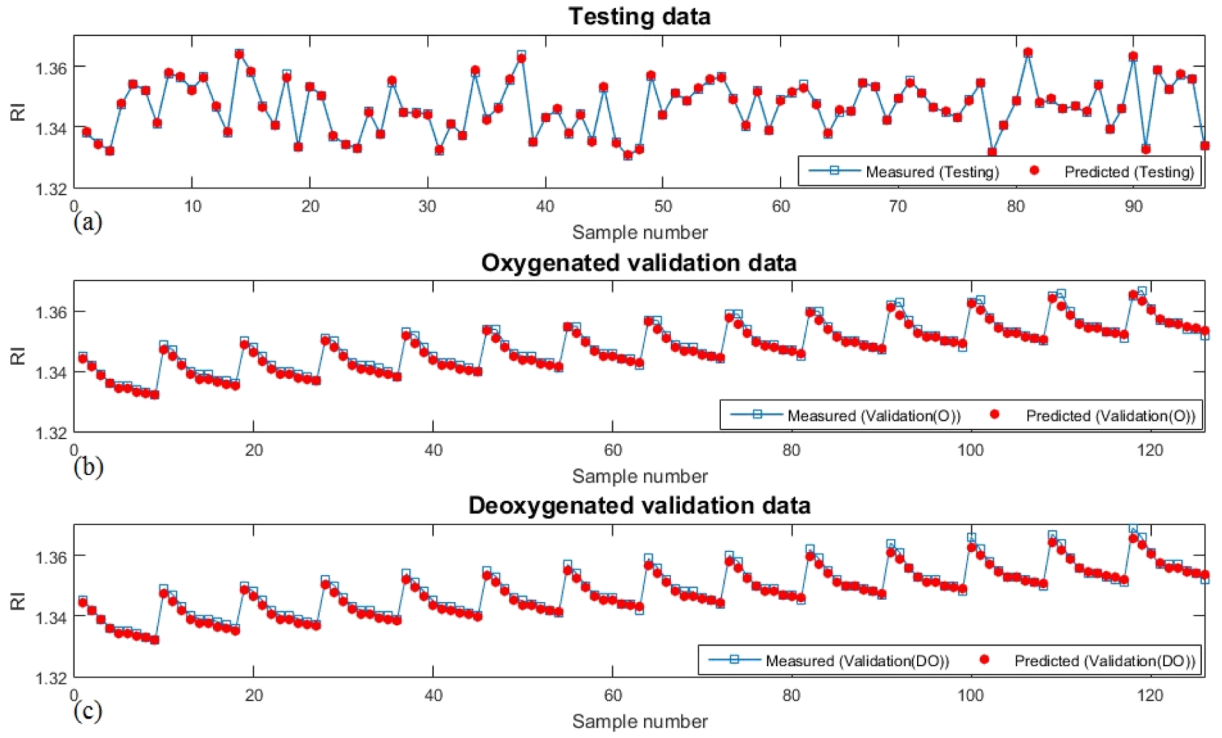


Figure 5. ACE-predicted RI in each sample number for (a) testing data, (b) validation (O), and (c) validation (DO). It is observed that estimated RI provides a good match with the measured values.

Table 3. Statistical Evaluation of Constructed Models

model	allocation	R2	MSE	AARE	SMAPE
ACE	training	0.9982336306	0.0000001065	0.0001908349	0.0000954193
	testing	0.9982064604	0.0000001291	0.0002052358	0.0001026208
	validation (O)	0.9717864827	0.0000016910	0.0007301520	0.0003652859
	validation (DO)	0.9743329323	0.0000015765	0.0007440273	0.0003722085
ONN	training	0.9952394041	0.0000002871	0.0003215436	0.0001607715
	testing	0.9941302810	0.0000004224	0.0003684964	0.0001842561
	validation (O)	0.9499491259	0.0000029999	0.0009546467	0.0004777142
	validation (DO)	0.9410967048	0.0000036180	0.0009860192	0.0004934819
OFIS	training	0.9941736595	0.0000003514	0.0003561574	0.0001780796
	testing	0.9888749901	0.0000008006	0.0004202847	0.0002102310
	validation (O)	0.9677079225	0.0000019355	0.0007854747	0.0003929896
	validation (DO)	0.9686669983	0.0000019245	0.0008053153	0.0004029079
OSVR	training	0.9944979014	0.0000003318	0.0003317496	0.0001658630
	testing	0.9946500649	0.0000003850	0.0003705688	0.0001852698
	validation (O)	0.9493983049	0.0000030329	0.0010064166	0.0005035914
	validation (DO)	0.9354325863	0.0000039659	0.0010239630	0.0005124898

Table 4. BA Regulation Parameters for Optimized Models

	value		
parameter	ONN	OFL	OSVR
number of variables for optimization	26	200	3
population size	100	400	10
maximum iteration	500	500	500
A_0	0.5	0.5	0.5
r_0	0.5	0.5	0.5
f_{min}	0	0	0
f_{max}	2	2	2

wavelength, optimal transformation of temperature, and sum of the input's optimal transformations, respectively. Figure 3 depicts calculated RI vs the sum of input's optimal transformations.

In this step, the equation which relates the value of RI and the summation value of the input's optimal transformations must be extracted. This task was done via the curve-fitting toolbox in the MATLAB environment. As shown in this graph, RI has a linear dependency on the sum of input's optimal transformations. The following equation, which relates RI to the sum of the input's optimal transformations, was derived through curve fitting

$$\mathrm{RI} = 0.00776533652401938S(\mathrm{Tr}) + 1.34598979050442 \quad (14)$$

Through this relationship, the equation that relates the value of RI to the summation of optimal transformations was

https://doi.org/10.1021/acsomega.2c00746
ACS Omega 2022, 7, 33769−33782

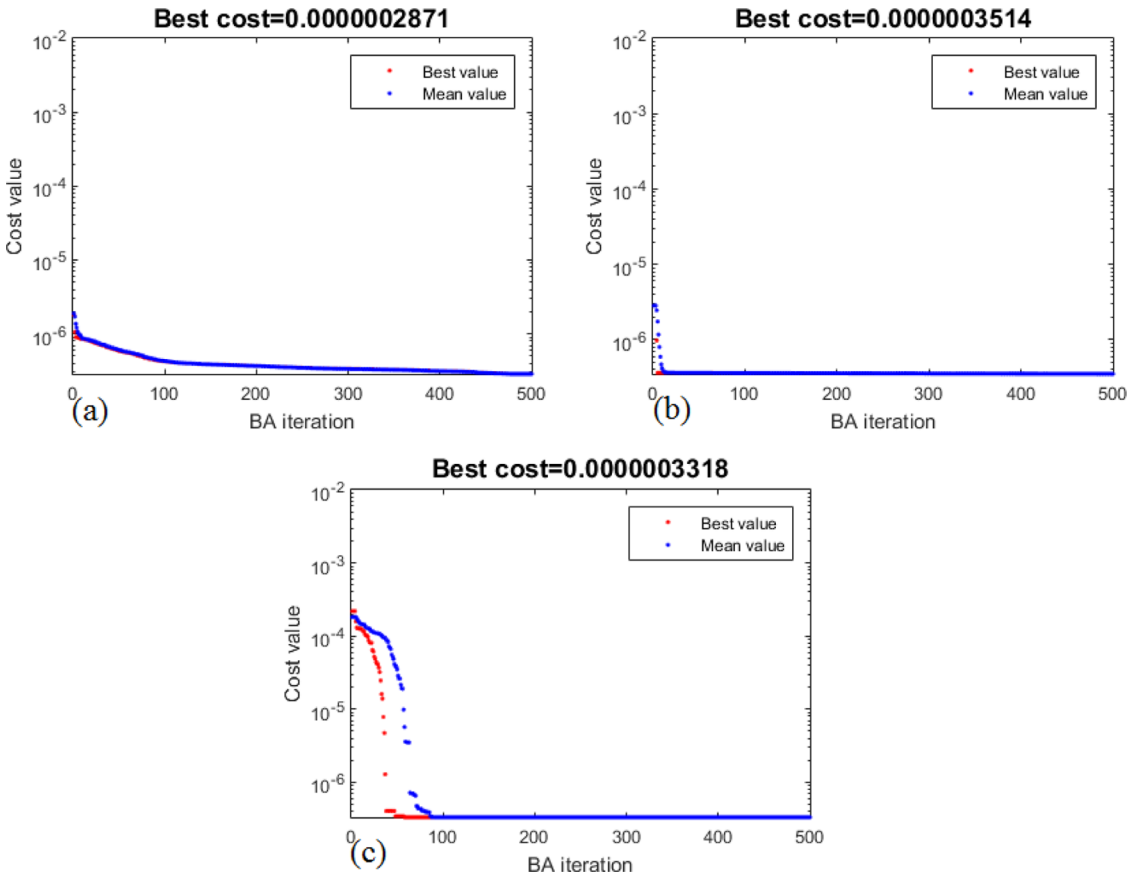


Figure 6. BA runs for optimization of (a) ONN, (b) OFIS, and (c) OSVR.

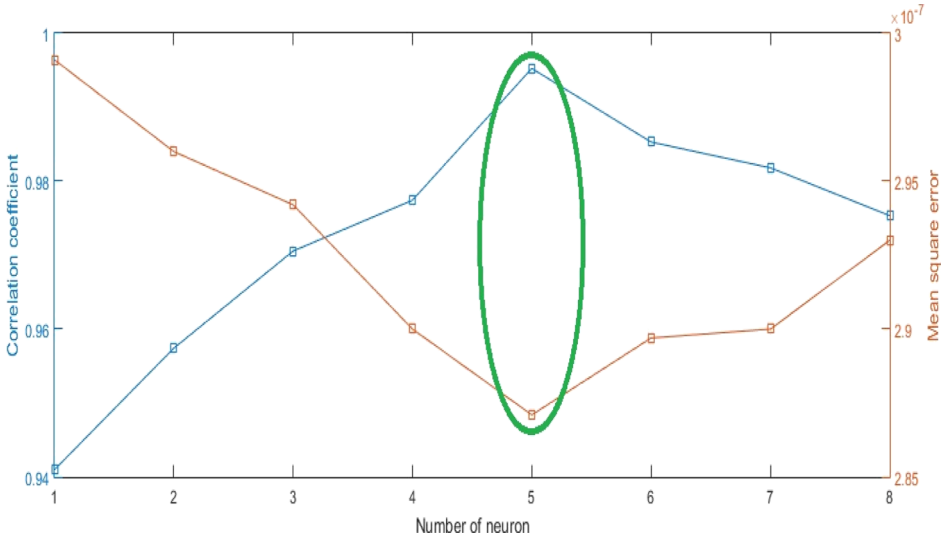


Figure 7. Statistical measurement for finding the best value for the number of neurons in the hidden layer. It is evident that 5 would be the best number since it makes the neural network have the maximum correlation of determination as well as minimum MSE.

determined. Substituting eqs 12 and 13 in eq 14, the correlation that relates RI to input parameters can be developed as follows

$$\mathrm{RI} = \alpha C + \beta T + \delta W + \lambda W^2 + \mu \quad (15)$$

The coefficients of the above equation are given in Table 2.

This equation represents how input parameters can be used for the prediction of hemoglobin's RI. In this stage, test data (84 data points) and validation data (126 data points of DO hemoglobin and 126 data points of O hemoglobin) were fed into eq 15, and the values of RI for these data points were obtained. Figure 4a demonstrates the cross-plot between actual RI and the RI values, which were estimated by the ACE approach for training, testing, and validation data points. As can be seen in this figure, ACE results have a perfect fit with observed data.

For a better comparison, in addition to Figure 4a, the correspondence of predicted RI with actual RI is assessed in each sample number (Figure 5). It is completely evident that calculated RI has an exact match with the experimental data.

https://doi.org/10.1021/acsomega.2c00746
ACS Omega 2022, 7, 33769−33782

Table 5. ONN Optimum Parameters Determined via BA

layer		weights			
	node	input 1	input 2	input 3	biases
hidden layer	node 1	2.3022	−0.1883	−0.3005	−2.1802
	node 2	−1.9735	−0.5759	−1.0198	0.4422
	node 3	0.3176	−0.7212	0.0201	−0.9432
	node 4	−0.4376	0.0769	0.0575	−0.1849
	node 5	1.4397	2.0913	−0.0002	2.1387

output layer	weights						
	node	node 1	node 2	node 3	node 4	node 5	bias
	node 1	0.0474	−0.0112	0.3732	−1.4738	−0.0642	0.0637

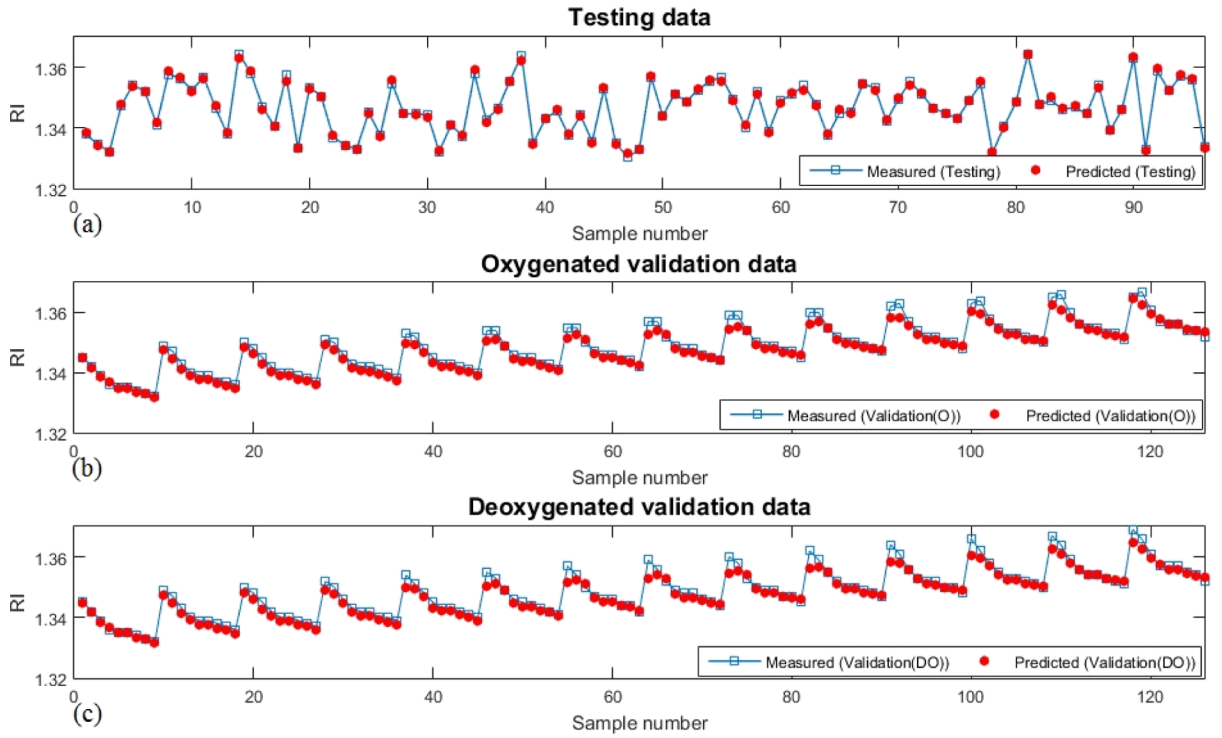


Figure 8. ONN-predicted RI in each sample number for (a) testing data, (b) validation (O), and (c) validation (DO), which confirms the validity of this method.

The results of this correlation were carefully assessed based on the statistical criteria defined in the previous part, and their results are listed in Table 3.

The output of statistical evaluation in all allocation of data (training, testing, validation (O), and validation (DO)) confirms that ACE has a very good performance in modeling hemoglobin's RI and therefore should be considered for quantitative estimation of RI.

4.2. ONN Results. In this part, a neural network combined with BA was used to build a model to determine the value of hemoglobin's RI. Training data points (384 data points) were used in the training stage to obtain the parameters of ONN. BA successfully keeps neural networks away from being trapped in local minima instead of global minima. Hence, ONN was developed based on the oversight to train the models. BA regulation parameters for ONN are given in Table 4.

In the constructed model, a hyperbolic tangent sigmoid was considered as a mapping function in hidden and output layers. Figure 6a depicts the process of BA running for optimizing the proposed neural network, leading to optimal calculation of weights and biases of the neural network.

One of the main issues in constructing a neural network model to implement in a regression task is to find out the optimum number of neurons in hidden layers. This number has a significant effect on the accuracy and complexity of developed models. Therefore, it is pretty crucial to precisely determine this number. Further analysis, which is depicted in Figure 7, was used for finding this optimum number, and based on this figure, it can be clearly deduced that 5 is the best value for the number of neurons that must be included in the hidden layer.

The optimum values of weights and biases of ONN are given in Table 5.

After constructing the model, test data and validation data were employed to assess the validity of the models. In Figure 4b, a comparison between the outputs of ONN and real values is presented in a cross-plot. Considering this figure, ONN has acceptable goodness-of-fit in predicting the RI. In Figure 8, the capability of ONN is displayed in each of the sample points. According to this figure, the ONN method is effective since observed and calculated data are very close. Moreover, statistical measurements of ONN outputs are listed in Table 3.

https://doi.org/10.1021/acsomega.2c00746
ACS Omega 2022, 7, 33769−33782

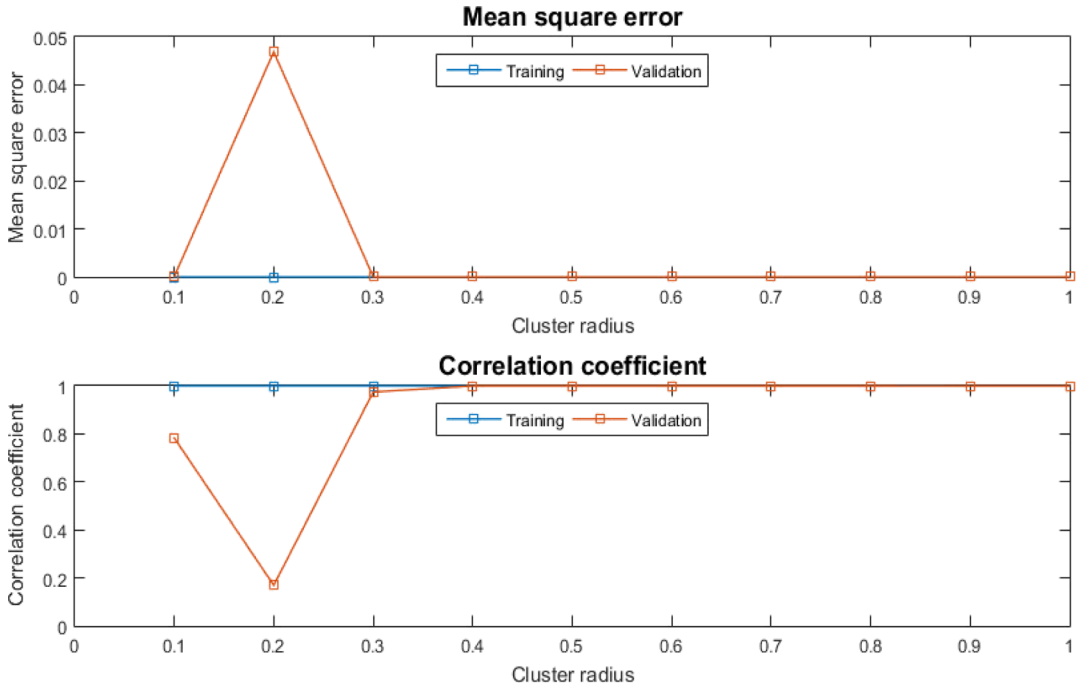


Figure 9. Statistical evaluation for extracting the cluster radius. Based on this figure, it can be found that 0.4 is the best value for the cluster radius to achieve the best models.

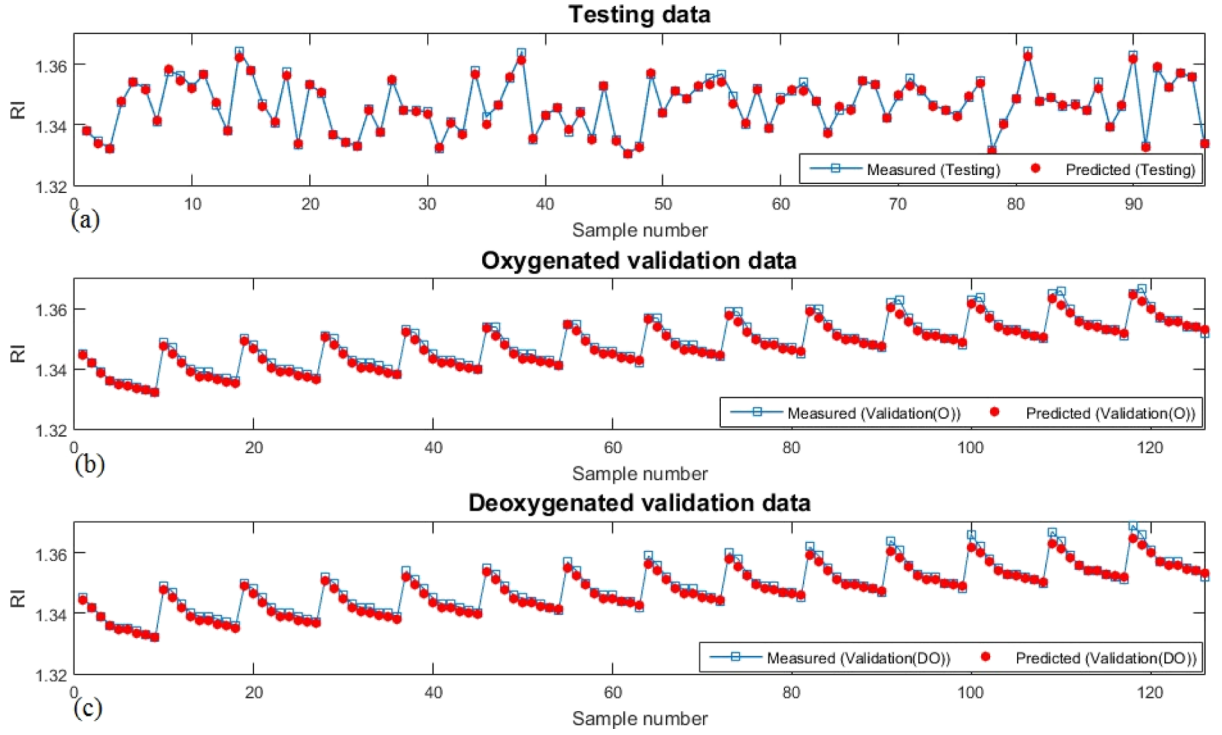


Figure 10. OFIS-predicted RI in each sample number for (a) testing data, (b) validation (O), and (c) validation (DO). For all portions of the data, outputs of OFIS highly match with measured values.

Table 6. Optimum Parameters of OSVR Determined by BA

parameter	value
gamma	7.72×10^{-5}
C	1.8937
epsilon	0.0015

4.3. OFIS Results. At first, training data points were adapted to learn the model and find its parameters. As mentioned before, improper regulation of membership functions reduces the efficiency of fuzzy inference systems. To avoid this problem, BA is integrated with a fuzzy inference system in order to reach the optimal values of its membership function. BA regulation parameters of OFIS are provided in Table 4. Results of running BA to disclose these suitable values are presented in Figure 6b.

https://doi.org/10.1021/acsomega.2c00746
ACS Omega 2022, 7, 33769−33782

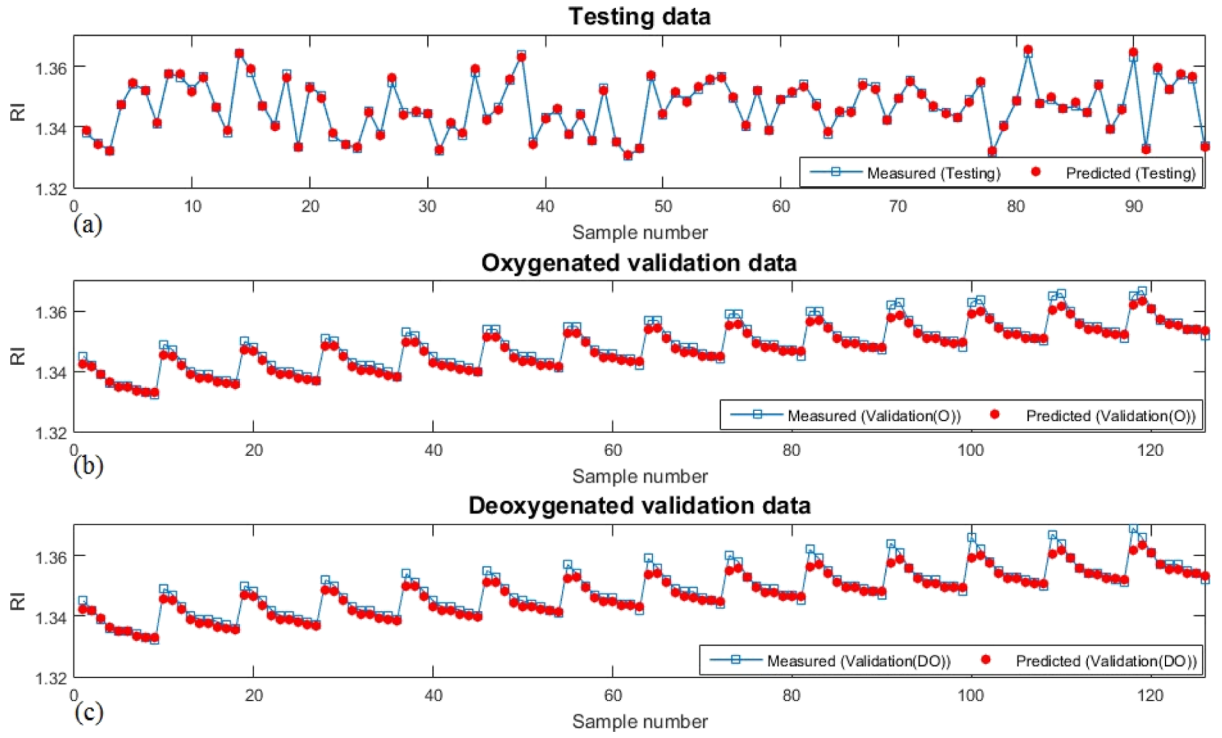


Figure 11. OSVR-predicted RI in each sample number for (a) testing data, (b) validation (O), and (c) validation (DO). It can be concluded that OSVR has a satisfactory capability in predicting the RI.

Table 7. Sensitivity Analysis of Constructed Models for Different Input Variables

input	ACE	ONN	OFIS	OSVR
concentration	66.51993	66.23167	66.34054	65.53819
wavelength	26.13098	26.62641	26.97092	26.90064
temperature	7.34910	7.14193	6.68853	7.56117

Moreover, statistical analysis for determining the best value of cluster radius is shown in Figure 9. This figure displays that 0.4 is the optimum value of the cluster radius.

After deciding the optimal value of the OFIS model, test data (96 data points) and validation data (252 data points) were imported in the constructed model, and the value of RI in these data points was calculated. After the RI is predicted and compared with real values (Figure 4c), it can be concluded that OFIS is a well-performing technique that can provide satisfactory results. Besides, in each sample point, the predicted and real values were compared (Figure 10). This figure illustrates capability of OFIS in correlating RI with its conditioning parameters.

Table 3 presents the statistical indexes of this model. According to this modeling scheme, OFIS is totally proper for the driving equation that proves the functional dependency between RI and its influencing parameters.

4.4. OSVR Results. To use OSVR in our modeling of RI, first, training data including 384 data points were employed to train the model. As mentioned in Section 2.4, there is a possibility that penalty parameters in support vector regression were adjusted inappropriately. To manage to successfully overcome this major deficiency in this study, support vector regression was merged with BA and the accuracy of the model was improved through extracting the optimum values of free parameters. BA regulation parameters for OSVR are summarized in Table 4. The BA performance is depicted in Figure 6c, and the optimal parameters of the support vector regression model are listed in Table 6.

After finding the optimal values of OSVR models, test data and validation data sets were included in constructing the model, and the values of RI were found. Figure 4d demonstrates the estimated RI values from OSVR vs the measured values of RI. In Figure 11, the comparison between predicted values and the observed values vs the sample numbers is represented. Based on these figures, it can be concluded that predicted RI values have satisfactory agreement with the observed values of RI.

Moreover, according to the statistical results presented in Table 3, it can be clearly concluded that this method is effective and viable for modeling RI.

4.5. Sensitivity Analysis. SA enables us to decide which input parameter would have the highest impact on the final outcome, which is the prediction of RI. For this purpose, an extremely simple structure was proposed based on Gandomi et al.'s[86] workflow. The following formulas are used to calculate the dependency of output on independent variables[86]

$$N_i = f_{\max}(x_i) - f_{\min}(x_i) \tag{16}$$

$$S_i = \frac{N_i}{\sum_{j=1}^{n} N_j} \times 100 \tag{17}$$

Here, $f_{\max}(x_i)$ and $f_{\min}(x_i)$ are the max and the min of the estimated values over the ith input domain, respectively, while mean values of other variables are replaced. Table 7 shows the results of SA for the proposed models. It explains that the highest influence on the RI of hemoglobin is from concentration (C).

4.6. Comparative Study. Table 3 presents the accuracy and general performance of the constructed models (ONN, OFIS, OSVR, and ACE) at a glance. This table evaluates the four models based on the assessed statistical criteria in section 4.1. As it can be seen in the table, results of ACE have smaller values of

https://doi.org/10.1021/acsomega.2c00746
ACS Omega 2022, 7, 33769−33782

AARE, SMAPE, and MSE and higher R^2 in comparison with the optimized models, which proves its superiority in estimation of RI. In addition, a visual comparison of performances of designed models is depicted in Figure 4. According to this figure, we understand that ACE's capability in estimation of RI is better compared to other AI models.

5. CONCLUSIONS

Obtaining a model to accurately determine the RI of hemoglobin is absolutely vital because this parameter plays a prominent role in medical diagnosis. To make quantitative formulations between the RI of hemoglobin and its influencing parameters, a statistical technique named ACE and three optimized models including ONN, OFIS, and OSVR were employed in this research. Results of the predictive models were appraised based on the statistical criteria. Moreover, SA was applied to measure the importance of each parameter in the estimation of RI. According to the simulation results of this study, the following conclusions can be achieved from this paper:

1 The efficiency of ACE in relating the RI of hemoglobin to input parameters was enhanced successfully. Moreover, all three optimized models' (ONN, OFIS, and OSVR) estimations of RI of hemoglobin were highly accurate when we compared predictions with observed values.
2 Capability of the BA in optimizing the weight and bias of a neural network could highly improve the model prediction performance. This was attained since the BA finds the optimum value of user-defined parameters of support vector regression with good accuracy.
3 The embedding of BA in the formulation of a fuzzy inference system led to optimal computation of membership functions with satisfactory precision to ultimately generate RI values with the least discrepancy with measured RI values.
4 Considering the evaluation of the models based on statistical criteria, the accuracy of ACE was found to be higher than those of ONN, OFIS, and OSVR. Finally, SA indicated that the important input parameters in the prediction of RI are concentration, wavelength, and temperature.
5 According to promising results of ACE, implementation of this methodology can conveniently eliminate the prohibitive cost of experimental measurement of hemoglobin's RI.3

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.2c00746.

Statistical evaluation of parameters that are employed for developing the models (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Mehdi Ostadhassan − *Department of Geology, Ferdowsi University of Mashhad, Mashhad 9177948974, Iran; Institute of Geosciences, Marine and Land Geomechanics and Geotectonics, Christian-Albrechts-Universität, Kiel 24118, Germany; Key Laboratory of Continental Shale Hydrocarbon Accumulation and Efficient Development, Ministry of Education, Northeast Petroleum University, Daqing 163318, China;* orcid.org/0000-0003-1285-5439; Email: mehdi.ostadhassan@ifg.kiel.de, mehdi.ostadhassan@nepu.edu.cn

Authors

Aida Alizamir − *Department of Pathology, School of Medicine, Hamadan University of Medical Science, Hamadan 6517838738, Iran*

Amin Gholami − *Reservoir Division, Iranian Offshore Oil Company, Tehran 1966653943, Iran*

Nader Bahrami − *Financial Transaction Department, Carsome Company, Petaling Jaya, Selangor 47800, Malaysia*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c00746

Author Contributions

All authors contributed to the study conception, design, and revisions. Conceptualization and coding: A.A. Data and methods: A.G. and N.B. Writing—original draft preparation: A.A., A.G., and N.B. Reviewing: M.O.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to respected reviewers who took their time to review this manuscript meticulously and provided us with constructive comments in two rounds. Moreover, we would like to thank respected editor for his editorial feedback and for giving us the opportunity to revise the manuscript and submit new versions for publication.

■ REFERENCES

(1) Faber, D. J.; Aalders, M. C. G.; Mik, E. G.; Hooper, B. A.; van Gemert, M. J. C.; van Leeuwen, T. G. Oxygen saturation-dependent absorption and scattering of blood. *Phys. Rev. Lett.* **2004**, *93*, 028102.

(2) Berkow, L. Factors affecting hemoglobin measurement. *J. Clin. Monit. Comput.* **2013**, *27*, 499−508.

(3) Park, Y.; Popescu, M.; Lykotrafitis, G.; Choi, G.; Feld, W.; Suresh, M. S.; Suresh, S. Refractive index maps and membrane dynamics of human red blood cells parasitized by Plasmodium falciparum. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 13730−13735.

(4) Serebrennikova, Y. M.; Huffman, D. E.; Garcia-Rubio, L. H. Characterization of red blood cells with multiwavelength transmission spectroscopy. *BioMed Res. Int.* **2015**, *2015*, 382641.

(5) Rinehart, M. T.; Park, H. S.; Walzer, K. A.; Chi, J. T. A.; Wax, A. Hemoglobin consumption by P. falciparum in individual erythrocytes imaged via quantitative phase spectroscopy. *Sci. Rep.* **2016**, *6*, 1−9.

(6) Wang, Z. Tissue refractive index as marker of disease. *J. Biomed. Opt.* **2011**, *16*, 116017.

(7) Jedrzejewska-Szczerska, M. Measurement of complex refractive index of human blood by low-coherence interferometry. *Eur. Phys. J. Spec. Top.* **2013**, *222*, 2367−2372.

(8) Jacques, S. L. Optical properties of biological tissues: A review. *Phys. Med. Biol.* **2013**, *58*, R37−61.

(9) Friebel, M.; Roggan, M.; Müller, A.; Meinke, I.; Roggan, A.; Gmbh, M. B. Determination of optical properties of human blood in the spectral range 250 to 1100 nm using Monte Carlo simulations with hematocrit-dependent effective scattering phase functions. *J. Biomed. Opt.* **2006**, *11*, 034021.

(10) Sardar, D. K.; Levy, L. B. Optical Properties of Whole Blood. *Laser Med. Sci.* **1998**, *13*, 106−111.

(11) Yahya, M.; Saghir, M. Z. Empirical modelling to predict the refractive index of human blood. *Phys. Med. Biol.* **2016**, *61*, 1405−1415.

(12) Tomari, R.; Zakaria, W. N. W.; Jamil, M. M. A.; Nor, F. M.; Fuad, N. F. N. Computer aided system for red blood cell classification in blood smear image. *Procedia Comput. Sci.* **2014**, *42*, 206−213.
(13) Kutlu, H.; Avci, E.; Özyurt, F. White blood cells detection and classification based on regional convolutional neural networks. *Med. Hypotheses* **2020**, *135*, 109472.
(14) Toğaçar, M.; Ergen, B.; Cömert, Z. Classification of white blood cells using deep features obtained from Convolutional Neural Network models based on the combination of feature selection methods. *Appl. Soft Comput.* **2020**, *97*, 106810.
(15) Patil, A. M.; Patil, M. D.; Birajdar, G. K. White blood cells image classification using deep learning with canonical correlation analysis. *IRBM* **2021**, *42*, 378−389.
(16) Girdhar, A.; Kapur, H.; Kumar, V. Classification of White blood cell using Convolution Neural Network. *Biomed. Signal Process Control* **2022**, *71*, 103156.
(17) Anita Davamani, K. A.; Rene Robin, C. R.; Doreen Robin, D. D.; Jani Anbarasi, L. J. Adaptive blood cell segmentation and hybrid Learning-based blood cell classification: A Meta-heuristic-based model. *Biomed. Signal Process Control* **2022**, *75*, 103570.
(18) Zecchin, C.; Facchinetti, A.; Sparacino, G.; Cobelli, C. Jump neural network for online short-time prediction of blood glucose from continuous monitoring sensors and meal information. *Comput. Methods Progr. Biomed.* **2014**, *113*, 144−152.
(19) Wang, S.; Yin, Y.; Cao, G.; Wei, B.; Zheng, Y.; Yang, G. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. *Neurocomputing* **2015**, *149*, 708−717.
(20) Soomro, T. A.; Afifi, A. J.; Gao, J.; Hellwich, O.; Zheng, L.; Paul, M. Strided fully convolutional neural network for boosting the sensitivity of retinal blood vessels segmentation. *Expert Syst. Appl.* **2019**, *134*, 36−52.
(21) Zhang, M.; Zhang, C.; Wu, X.; Cao, X.; Young, G. S.; Chen, H.; Xu, X. A neural network approach to segment brain blood vessels in digital subtraction angiography. *Comput. Methods Progr. Biomed.* **2020**, *185*, 105159.
(22) Tchinda, B. S.; Tchiotsop, D.; Noubom, M.; Louis-Dorr, V.; Wolf, D. Retinal blood vessels segmentation using classical edge detection filters and the neural network. *Inform. Med. Unlocked* **2021**, *23*, 100521.
(23) Gegundez-Arias, M. E.; Marin-Santos, D.; Perez-Borrero, I.; Vasallo-Vazquez, M. J. A new deep learning method for blood vessel segmentation in retinal images based on convolutional kernels and modified U-Net model. *Comput. Methods Progr. Biomed.* **2021**, *205*, 106081.
(24) Deng, X.; Ye, J. A retinal blood vessel segmentation based on improved D-MNet and pulse-coupled neural network. *Biomed. Signal Process Control* **2022**, *73*, 103467.
(25) Zhang, Y.; He, M.; Chen, Z.; Hu, K.; Li, X.; Gao, X. Bridge-Net: Context-involved U-net with patch-based loss weight mapping for retinal blood vessel segmentation. *Expert Syst. Appl.* **2022**, *195*, 116526.
(26) Xu, Z.; Liu, J.; Chen, X.; Wang, Y.; Zhao, Z. Continuous blood pressure estimation based on multiple parameters from eletrocardiogram and photoplethysmogram by Back-propagation neural network. *Comput. Ind.* **2017**, *89*, 50−59.
(27) Senturk, U.; Polat, K.; Yucedag, I. A non-invasive continuous cuffless blood pressure estimation using dynamic Recurrent Neural Networks. *Appl. Acoust.* **2020**, *170*, 107534.
(28) Esmaelpoor, J.; Moradi, M. H.; Kadkhodamohammadi, A. A multistage deep neural network model for blood pressure estimation using photoplethysmogram signals. *Comput. Biol. Med.* **2020**, *120*, 103719.
(29) Baker, S.; Xiang, W.; Atkinson, I. A hybrid neural network for continuous and non-invasive estimation of blood pressure from raw electrocardiogram and photoplethysmogram waveforms. *Comput. Methods Progr. Biomed.* **2021**, *207*, 106191.
(30) Qiu, Y.; Liu, D.; Yang, G.; Qi, D.; Lu, Y.; He, Q.; Qian, J.; Li, X.; Cao, Y.; Shuai, J. Cuffless blood pressure estimation based on composite neural network and graphics information. *Biomed. Signal Process Control* **2021**, *70*, 103001.
(31) Cheng, J.; Xu, Y.; Song, R.; Liu, Y.; Li, C.; Chen, X. Prediction of arterial blood pressure waveforms from photoplethysmogram signals via fully convolutional neural networks. *Comput. Biol. Med.* **2021**, *138*, 104877.
(32) Ali, J. B.; Hamdi, T.; Fnaiech, N.; Di Costanzo, V.; Fnaiech, F.; Ginoux, J. M. Continuous blood glucose level prediction of type 1 diabetes based on artificial neural network. *Biocybern. Biomed. Eng.* **2018**, *38*, 828−840.
(33) D'Antoni, F.; Merone, M.; Piemonte, V.; Iannello, G.; Soda, P. Auto-Regressive Time Delayed jump neural network for blood glucose levels forecasting. *Knowl. Base Syst.* **2020**, *203*, 106134.
(34) Alfian, G.; Syafrudin, M.; Anshari, M.; Benes, F.; Atmaji, F. T. D.; Fahrurrozi, I.; Hidayatullah, J.; Rhee, J. Blood glucose prediction model for type 1 diabetes based on artificial neural network with time-domain features. *Biocybern. Biomed. Eng.* **2020**, *40*, 1586−1599.
(35) Dudukcu, H. V.; Taskiran, M.; Yildirim, T. Blood glucose prediction with deep neural networks using weighted decision level fusion. *Biocybern. Biomed. Eng.* **2021**, *41*, 1208−1223.
(36) Zhang, M.; Flores, K. B.; Tran, H. T. Deep learning and regression approaches to forecasting blood glucose levels for type 1 diabetes. *Biomed. Signal Process Control* **2021**, *69*, 102923.
(37) Alade, I. O.; Bagudu, A.; Oyehan, T. A.; Rahman, M. A.; Saleh, T. A.; Olatunji, S. Estimating the refractive index of oxygenated and deoxygenated hemoglobin using genetic algorithm - support vector regression model. *Comput. Methods Progr. Biomed.* **2018**, *163*, 135−142.
(38) Oyehan, T. A.; Alade, I. O.; Bagudu, A.; Sulaiman, K. O.; Olatunji, S. O.; Saleh, T. A. Predicting of the refractive index of haemoglobin using the Hybrid GA-SVR approach. *Comput. Biol. Med.* **2018**, *98*, 85−92.
(39) Breiman, L.; Friedman, J. H. Estimating optimal transformations for multiple regression and correlation. *J. Am. Stat. Assoc.* **1985**, *80*, 580−598.
(40) Gholami, A.; Moradi, S.; Asoodeh, M.; Bagheripour, P.; Vaezzadeh-Asadi, M. Asphaltene precipitation modeling through ACE reaping of scaling equations. *Sci. China Chem.* **2014**, *57*, 1774−1780.
(41) Gholami, A.; Asoodeh, M.; Bagheripour, P. How committee machine with SVR and ACE estimates bubble point pressure of crudes. *Fluid Phase Equilib.* **2014**, *382*, 139−149.
(42) Gholami, A. Oil Formation Volume Factor Determination Through a Fused Intelligence. *Acta Geophys.* **2016**, *64*, 2510−2529.
(43) Gholami, A.; Mohammadzadeh, O.; Kord, S.; Moradi, S.; Dabir, B. Improving the estimation accuracy of titration-based asphaltene precipitation through power-law committee machine (PLCM) model with alternating conditional expectation (ACE) and support vector regression (SVR) elements. *J. Pet. Explor. Prod. Technol.* **2016**, *6*, 265−277.
(44) Golden, R. M. *Mathematical Methods for Neural Network Analysis and Design*; MIT Press: Cambridge, MA, USA, 1996.
(45) Yadav, A.; Chatterjee, S.; Equeenuddin, S. M. Suspended sediment yield estimation using genetic algorithm-based artificial intelligence models: case study of Mahanadi River, India. *Hydrol. Sci. J.* **2018**, *63*, 1162−1182.
(46) Noori, R.; Deng, Z.; Kiaghadi, A.; Kachoosangi, F. T. How reliable are ANN, ANFIS, and SVM techniques for predicting longitudinal dispersion coefficient in natural rivers? *J. Hydraul. Eng.* **2016**, *142*, 04015039.
(47) Khan, M. Y. A.; Tian, F.; Hasan, F.; Chakrapani, G. J. Artificial neural network simulation for prediction of suspended sediment concentration in the River Ramganga, Ganges Basin, India. *Int. J. Sediment Res.* **2019**, *34*, 95−107.
(48) Afan, H. A.; El-shafie, A.; Mohtar, W. H. M. W.; Yaseen, Z. M. Past, present and prospect of an Artificial Intelligence (AI) based model for sediment transport prediction. *J. Hydrol.* **2016**, *541*, 902−913.
(49) Haddadchi, A.; Movahedi, N.; Vahidi, E.; Omid, M. H.; Dehghani, A. A. Evaluation of suspended load transport rate using transport formulas and artificial neural network models (Case study: Chelchay Catchment). *J. Hydrodyn.* **2013**, *25*, 459−470.

https://doi.org/10.1021/acsomega.2c00746
ACS Omega 2022, 7, 33769−33782

(50) Asoodeh, M.; Gholami, A.; Bagheripour, P. Oil-CO2MMP Determination in Competition of Neural Network, Support Vector Regression, and Committee Machine. *J. Dispersion Sci. Technol.* **2014**, *35*, 564−571.
(51) Gholami, A.; Ansari, H. R.; Ahmadi, S. Combining of intelligent models through committee machine for estimation of wax deposition. *J. Chin. Chem. Soc.* **2018**, *65*, 925−931.
(52) Kakaei Lafdani, E. K.; Moghaddam Nia, A. M.; Ahmadi, A. Daily suspended sediment load prediction using artificial neural networks and support vector machines. *J. Hydrol.* **2013**, *478*, 50−62.
(53) Gholami, A.; Afshar, M.; Bagheripour, P.; Asoodeh, M.; Vaezzadeh-Asadi, M. Smart correlation of compositional data to saturation pressure. *J. Nat. Gas Sci. Eng.* **2015**, *22*, 661−669.
(54) Zargar, G.; Bagheripour, P.; Asoodeh, M.; Gholami, A. Oil-CO2minimum miscible pressure (MMP) determination using a stimulated smart approach. *Can. J. Chem. Eng.* **2015**, *93*, 1730−1735.
(55) Gholami, M.; Bodaghi, A. A robust approach through combining optimized neural network and optimized support vector regression for modeling deformation modulus of rock masses. *Model. Earth Syst. Environ.* **2017**, *3*, 1−9.
(56) Gholami, A.; Seyedali, S. M.; Ansari, H. R. Estimation of shear wave velocity from post-stack seismic data through committee machine with cuckoo search optimized intelligence models. *J. Pet. Sci. Eng.* **2020**, *189*, 106939.
(57) Afshar, M.; Gholami, A.; Asoodeh, M. Genetic optimization of neural network and fuzzy logic for oil bubble point pressure modeling. *Korean J. Chem. Eng.* **2014**, *31*, 496−502.
(58) Asoodeh, M.; Gholami, A.; Bagheripour, P. Asphaltene precipitation of titration data modeling through committee machine with stochastically optimized fuzzy logic and optimized neural network. *Fluid Phase Equilib.* **2014**, *364*, 67−74.
(59) Gholami, A.; Asoodeh, M.; Bagheripour, P. Smart determination of difference index for asphaltene stability evaluation. *J. Dispersion Sci. Technol.* **2014**, *35*, 572−576.
(60) Asoodeh, M.; Bagheripour, P.; Gholami, A. NMR parameters determination through ACE committee machine with genetic implanted fuzzy logic and genetic implanted neural network. *Acta Geophys.* **2015**, *63*, 735−760.
(61) Zadeh, L. A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338−353.
(62) Gholami, A.; Asoodeh, M.; Bagheripour, P. Fuzzy assessment of asphaltene stability in crude oils. *J. Dispersion Sci. Technol.* **2014**, *35*, 556−563.
(63) Ahmadi, S.; Amiribakhtiar, M. S.; Gholami, A.; Bahrami, N. Upgrading fuzzy logic by GA-PS to determine asphaltene stability in crude oil. *Egypt. J. Pet.* **2017**, *26*, 505−510.
(64) Zargar, G.; Gholami, A.; Asoodeh, M.; Bagheripour, P. PSO-fuzzy eliminates deficiency of neuro-fuzzy in assessment of asphaltene stability. *Indian J. Chem. Technol.* **2016**, *22*, 135−140.
(65) Vapnik, V. *The Nature of Statistical Learning Theory*; Springer: New York, 1995.
(66) Bagheripour, P.; Gholami, A.; Asoodeh, M. Support vector regression between PVT data and bubble point pressure. *J. Pet. Explor. Prod. Technol.* **2015**, *5*, 227−231.
(67) Raghavendra N, S.; Deka, P. C. Support vector machine applications in the field of hydrology: A review. *Appl. Soft Comput.* **2014**, *19*, 372−386.
(68) Na'imi, S. R.; Gholami, A.; Asoodeh, M. Prediction of crude oil asphaltene precipitation using support vector regression. *J. Dispersion Sci. Technol.* **2014**, *35*, 518−523.
(69) Bagheripour, P.; Gholami, A.; Asoodeh, M.; Vaezzadeh-Asadi, M. Support vector regression based determination of shear wave velocity. *J. Pet. Sci. Eng.* **2015**, *125*, 95−99.
(70) Bodaghi, A.; Ansari, H. R.; Gholami, M. Optimized support vector regression for drillingrate of penetration estimation. *Open Geosci.* **2015**, *7*, 870−879.
(71) Fei, S. W.; Wang, M. J.; Miao, Y. B.; Tu, J.; Liu, C. L. Particle swarm optimization-based support vector machine for forecasting dissolved gases content in power transformer oil. *Energy Convers. Manage.* **2009**, *50*, 1604−1609.
(72) Fei, S. W.; Liu, C. L.; Miao, Y. B. Support vector machine with genetic algorithm for forecasting of key-gas ratios in oil-immersed transformer. *Expert Syst. Appl.* **2009**, *36*, 6326−6331.
(73) Gholami, M.; Bodaghi, A. Fusing of optimized intelligence models by virtue of committee machine for estimation of the residual shear strength of clay. *Model. Earth Syst. Environ.* **2016**, *2*, 43−52.
(74) Ansari, H. R.; Gholami, A. An improved support vector regression model for estimation of saturation pressure of crude oils. *Fluid Phase Equilib.* **2015**, *402*, 124−132.
(75) Ansari, H. R.; Gholami, A. Robust method based on optimized support vector regression for modeling of asphaltene precipitation. *J. Pet. Sci. Eng.* **2015**, *135*, 201−205.
(76) Bozorg-Haddad, O.; Soleimani, S.; Loáiciga, H. A. Modeling Water-Quality Parameters Using Genetic Algorithm−Least Squares Support Vector Regression and Genetic Programming. *J. Environ. Eng.* **2017**, *143*, 04017021.
(77) Fattahi, H.; Gholami, A.; Amiribakhtiar, M. S.; Moradi, S. Estimation of asphaltene precipitation from titration data: a hybrid support vector regression with harmony search. *Neural Comput. Appl.* **2015**, *26*, 789−798.
(78) Gholami, A.; Ansari, H. R. Estimation of porosity from seismic attributes using a committee model with bat-inspired optimization algorithm. *J. Pet. Sci. Eng.* **2017**, *152*, 238−249.
(79) Gholami, A.; Ansari, H. R.; Hosseini, S. Prediction of crude oil refractive index through optimized support vector regression: a competition between optimization techniques. *J. Pet. Explor. Prod. Technol.* **2017**, *7*, 195−204.
(80) Liao, R.; Zheng, H.; Grzybowski, S.; Yang, L. Particle swarm optimization-least squares support vector regression based forecasting model on dissolved gases in oil-filled power transformers. *Electr. Power Syst. Res.* **2011**, *81*, 2074−2080.
(81) Meysam Mousavi, S. M.; Tavakkoli-Moghaddam, R.; Vahdani, B.; Hashemi, H.; Sanjari, M. J. A new support vector model-based imperialist competitive algorithm for time estimation in new product development projects. *Comput. Integr. Manuf.* **2013**, *29*, 157−168.
(82) Rahgoshay, M.; Feiznia, S.; Arian, M.; Hashemi, S. A. A. Modeling daily suspended sediment load using improved support vector machine model and genetic algorithm. *Environ. Sci. Pollut. Res.* **2018**, *25*, 35693−35706.
(83) Roushangar, K.; Koosheh, A. Evaluation of GA-SVR method for modeling bed load transport in gravel-bed rivers. *J. Hydrol.* **2015**, *527*, 1142−1152.
(84) Yang, X. S. Bat algorithm and cuckoo search: a tutorial. 2013. In *Artificial intelligence, evolutionary computing and metaheuristics*; Springer: Berlin, Heidelberg, pp 421−434. DOI: 10.1007/978-3-642-29694-9_17.
(85) Zhernovaya, O.; Sydoruk, O.; Tuchin, V.; Douplik, A. The refractive index of human hemoglobin in the visible range. *Phys. Med. Biol.* **2011**, *56*, 4013−4021.
(86) Gandomi, A. H.; Yun, G. J.; Alavi, A. H. An evolutionary approach for modeling of shear strength of RC deep beams. *Mater. Struct.* **2013**, *46*, 2109−2119.

https://doi.org/10.1021/acsomega.2c00746
ACS Omega 2022, 7, 33769−33782