

Taylor & Francis
Taylor & Francis Group

Computer Methods in Biomechanics and Biomedical Engineering

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gcmb20

Dragonfly algorithm–support vector machine approach for prediction the optical properties of blood

Faiza Omari, Latifa Khaouane, Maamar Laidi, Abdellah Ibrir, Mohamed Roubehie Fissa, Mohamed Hentabli & Salah Hanini

To cite this article: Faiza Omari, Latifa Khaouane, Maamar Laidi, Abdellah Ibrir, Mohamed Roubehie Fissa, Mohamed Hentabli & Salah Hanini (2023): Dragonfly algorithm–support vector machine approach for prediction the optical properties of blood, Computer Methods in Biomechanics and Biomedical Engineering, DOI: 10.1080/10255842.2023.2228957

To link to this article: https://doi.org/10.1080/10255842.2023.2228957

Published online: 28 Jun 2023.

Submit your article to this journal

Article views: 24

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=gcmb20

COMPUTER METHODS IN BIOMECHANICS AND BIOMEDICAL ENGINEERING
https://doi.org/10.1080/10255842.2023.2228957

Taylor & Francis
Taylor & Francis Group

Check for updates

Dragonfly algorithm–support vector machine approach for prediction the optical properties of blood

Faiza Omari[a], Latifa Khaouane[a], Maamar Laidi[a], Abdellah Ibrir[a,b], Mohamed Roubehie Fissa[a], Mohamed Hentabli[a,c] and Salah Hanini[a]

[a]Laboratory of Biomaterials and Transport Phenomena (LBMTP), Yahia Fares University, Medea, Algeria; [b]Materials and Environment Laboratory (LME), Faculty of Technology, Yahia Fares University, Medea, Algeria; [c]Quality Control Laboratory, SAIDAL Complex of Medea, Medea, Algeria

ABSTRACT

Knowledge of the optical properties of blood plays important role in medical diagnostics and therapeutic applications in laser medicine. In this paper, we present a very rapid and accurate artificial intelligent approach using Dragonfly Algorithm/Support Vector Machine models to estimate the optical properties of blood, specifically the absorption coefficient, and the scattering coefficient using key parameters such as wavelength (nm), hematocrit percentage (%), and saturation of oxygen (%), in building very highly accurate Dragonfly Algorithm-Support Vector Regression models (DA-SVR). 1000 training and testing sets were selected in the wavelength range of 250-1200 nm and the hematocrit of 0-100%. The performance of the proposed method is characterized by high accuracy indicated in the correlation coefficients (*R*) of 0.9994 and 0.9957 for absorption and scattering coefficients, respectively. In addition, the root mean squared error values (*RMSE*) of 0.972 and 2.9193, as well as low mean absolute error values (*MAE*) of 0.2173 and 0.2423, this result showed a strong match with the experimental data. The models can be used to accurately predict the absorption and scattering coefficients of blood, and provide a reliable reference for future studies on the optical properties of human blood.

ARTICLE HISTORY
Received 21 March 2023
Accepted 19 June 2023

KEYWORDS
Absorption coefficient; scattering coefficient; human blood; dragonfly algorithm; support vector machine

1. Introduction

The optical properties of blood, are important parameters in biomedical optics, and have been extensively studied in the literature. Hemoglobin (HB), the main component of blood, allows the transport of oxygen from the lungs to the human tissues and plays a crucial role in the optical properties of blood. The absorption and scattering properties of hemoglobin depend on numerous physiological and biochemical parameters, including its concentration, hematocrit HCT (the volume fraction of red blood cells in blood), oxygen saturation (the percentage of hemoglobin that is bound to oxygen), and other parameters. (Faber et al. 2004; Friebel and Meinke 2005; Friebel et al. 2009).

Various experimental techniques have been used to measure the optical properties of blood, such as spectrophotometry, spectroradiometry, and reflectance/transmission measurements. These techniques typically involve measuring the intensity of light before and after passing through or interacting with blood, and then using mathematical models to extract the absorption and scattering properties of blood.

Friebel and Meinke (2005) determined the absorption coefficient of blood by transmittance measurements that have given by Beer-Lambert's law: $T = \exp(-\mu_a d)$ where d is the thickness of the sample. The integrating sphere measurements and inverse Monte Carlo simulations have been made by Friebel et al. (2006) for highly diluted blood. Roggan et al. (1999) and Yaroslavsky et al. (1996) measured μ_a and μ_s of human blood, their results showing that blood absorption followed the absorption behavior of Hb and water. Some studies performed the scattering and absorption cross sections using Mie theory, and Rayleigh–Gans approximation (Hammer et al. 1998).

The effect of optical properties of platelets and plasma on the optical behavior of whole blood was investigated by Meinke et al. (2007a, b), their results showed that the optical properties of red blood cells (RBCs) are greater compared to the other blood components; it is important to take this into account

CONTACT Faiza Omari omari.faiza@gmail.com
© 2023 Informa UK Limited, trading as Taylor & Francis Group

when making statements concerning the optical parameters of whole blood.

Until the present, it is regrettable that only a few theoretical and experimental studies on the optical properties of blood have reliable results as evidenced by the discrepancies in the findings reported in the literature, these might be caused by the various experimental techniques and measurement devices used in these works. (Zhernovaya et al. 2011; Yahya and Saghir 2016). In addition, there are several limitations and challenges in measuring the optical properties of blood. One major challenge is the potential issues related to aggregation of red blood cells at high concentrations, which can alter the scattering properties of blood and affect the accuracy of measurement results. Another challenge is the absorption of light by other chromophores in blood, such as water, lipids, and other proteins, which can interfere with measurement of optical properties of blood is also affected by the wavelength range of interest. As well, physiological and biochemical factors, such as pH, temperature that can cause changes in optical behavior of blood, (as higher temperature leads to damage to the shape of the red blood cells RBCs from disc-shaped to spherical, denaturation, and aggregation of blood), and the presence of other chemicals, can also affect the optical properties of blood and introduce variability in the measurement results. (Nilsson et al. 1997; Mordon et al. 2005; Steinke and Shepherd 1992). Higher hemoglobin concentration, hematocrit, and oxygen saturation generally result in higher absorption of light by blood, particularly in the visible range. In the near infrared range, the absorption is lower, but scattering is still significant, which can also impact the accuracy of measurement results (Nilsson et al. 1997; Mordon et al. 2005; Fu et al. 2011).

The detailed understanding of the interaction between light and blood is pivotal in various medical applications, including routine medical diagnostics, hematology, cardiac surgery, and therapeutic applications in laser medicine (Sardar and Levy 1998; Sardar et al. 2007; Jacques 2013). Additionally, optical techniques such as optical tomography, optical biopsy, diaphanoscopic, laser-induced thermotherapy, and others, heavily rely on the optical properties of blood to assess light distribution in blood-perfused tissues (Roggan et al. 1999; Jin et al. 2006). The absorption (μ_a) and scattering (μ_s) properties of hemoglobin are particularly valuable in these techniques as they provide insights into biophysical cell parameters and shed light on blood cell diseases, including cancer (Prince and Malarvizhi 2010). However, despite their significance, there is limited scientific research available in the literature on the absorption and scattering properties of blood, in contrast to the widely studied and estimated the refractive index using various machine learning approaches. (Yahya and Saghir 2016; Oyehan et al. 2018; Alade et al. 2018; Oduncuoğlu 2021; Rakhshani and Mansouri-Birjandi 2017; Alizamir et al. 2022).

Additionally, some of these techniques do not consider the most important factors that influence the optical properties, key examples include the oxygen saturation (SO2), and hematocrit (HCT). To fill this gap, it is necessary to develop new and accurately strategy that can be reliably used to estimate the optical properties of blood.

According to what addressed before, machine learning (ML) approaches have demonstrated their predictive power in solving challenging issues in diverse fields including chemical and biomedical engineering, petroleum industry, water purification, and environment protection (Al-Jamimi et al.2019; Alade et al.2019; Wang et al. 2012; Al-Jamimi et al. 2018; Saleh 2021).

The aim of this work is to develop Dragonfly Algorithm/Support Vector Regression (DA-SVR) models to predict the optical properties of human blood, specifically, the absorption coefficient (μ_a) and the scattering coefficient (μ_s), using the previous physiological and biochemical parameters for the first time, which makes the models can be easily used in the full range of wavelengths with flexibility, and have the potential to be used for the real world applications in various fields, ranging from non-invasive blood monitoring, medical diagnostics (anemia, diabetes, and cancer) (Zhernovaya et al.2007; Prince and Malarvizhi 2010) which are associated with changes in blood optical properties, drug development, biomedical research, to clinical decision support, and can contribute to advancements in healthcare and medical technologies.

2. Materials and methods

2.1. Description of the support Vector machine regression

The SVM is one of the powerful machine learning techniques, based on statistical learning theory (SLT) which is proposed by Vapnik and Corte (Vladimir Vapnik 1995; Boser et al. 1992) it is a reliable algorithm for the classification and regression of linear and non-linear problems (Drucker et al. 1997) in the principle of structural risk minimization (SRM)

instead of empirical risk minimization based Neural networks (Byvatov et al. 2003), this improves the ability of SVR in prediction even when the dataset and their descriptors are few (Benimam et al. 2020; Ibrir et al. 2021; Hariharan et al. 2018; Hung and Chiu 2017).

Given a training dataset (x_i, y_i) with $i = 1, 2 \ldots n$, x, and y represent the model inputs and output, and n is the number of data points.

The objective of support vector regression is to find a regression function to produce the target $\hat{y}$ it can be estimated by

$$\hat{y} = \omega^{T}.\emptyset(x) + b \tag{1}$$

where $\emptyset$ represents the nonlinear function, ω is a weight vector that can be adjusted, and b is the bias which is the scaler. To find the minimum error or deviation ε, the smallest weight vector must be found by minimizing the $\|\omega\|^2$ and measuring the same permissible deviation on non-negative slack variable ε and ε^* representing upper and lower learning errors, this leads to the following expression:

$$\text{Minimise}: \frac{1}{2} \quad \|\omega\|^2 + C \sum_{i=1}^{N} (\varepsilon_i + \varepsilon_i) \tag{2}$$

$$\text{subject to} \begin{cases} y_i - \hat{y}_i = y_i - (\omega^T.\emptyset(x) + b) \leq \varepsilon + \varepsilon_i \\ \hat{y}_i - y_i = (\omega^T.\emptyset(x) + b) - y_i \leq \varepsilon + \varepsilon_i \end{cases}$$

The constant $C > 0$ is called penalty parameter or regularized parameter that measures the trade-off between the complexity of the model and empirical error (He et al. 2014), according to the method of Lagrange multiples, (L) the above optimization problem in Equation (2) can be solved to transform it into dual variables (ω; C; ε; ε^*):

Min

$$\sum_{i=1}^{N} y_i \left(\alpha_i - \alpha_i'\right) - \varepsilon \sum_{i=1}^{N} (\alpha_i - \alpha_i') - \frac{1}{2} \sum_{i=1}^{N} \sum_{i=1}^{N} (\alpha_i - \alpha_i') \left(\alpha_j - \alpha_j'\right) \emptyset(x)^{T}.\emptyset(x) \tag{3}$$

$$\text{Subject to} \sum_{i=1}^{N} (\alpha_i - \alpha_i') = 0 \quad (i = 1, 2 \ldots, N)$$

$$0 \leq (\alpha_i - \alpha_i') \leq C$$

where α and α' are Lagrange multipliers. Which are solved by a standard quadratic programming algorithm, then the regression function becomes

$$f(x) = \sum_{i=1}^{N} (\alpha_I - \alpha_I') k(x_i, x_j) + b \tag{4}$$

where (α_i - α_i') $\neq 0$, α_i and α_i' are support vectors, N is the number of support vectors, b is the bias, and the kernel function k, which facilitates solving nonlinear problems linearly, it is a measure of similarities between input vectors in feature space (Smola et al. 1998)

$$k(x_i, x_j) = \emptyset(x_i)^{T} \emptyset(x_j) \tag{5}$$

There are three different kernel functions that have a significant effect on the SVM's effectiveness (Amari and Wu 1999) which are;

- Linear, $k(x_i, x_j) = (x_i, x_j)$ (6)
- Polynomial, $k(x_i, x_j) = [\gamma (x_i, x_j) + 1]^{d}$, $\gamma > 0$ (7)
- Gaussian or radial basis function (RBF)

$$k(x_i, x_j) = \exp(-\gamma . |x_i, x_j|^2) \tag{8}$$

where d, is the degree of the polynomial, and γ is the kernel parameter

2.2. Dragonfly algorithm

Recently, Meta-heuristic algorithms have been used to optimize the SVM's parameters, such as genetic algorithm (GA) (Xie et al. 2019), particle swarm optimization (PSO) (Li et al. 2020), and the dragonfly algorithm (DA) (Moussa et al. 2022), which is proposed by Mirjalili (2016) based on the natural swarming behaviors of dragonflies, both static and dynamic (Chatterjee et al. 2022).

The Dragonfly Algorithm is based on the collective behavior and movement patterns of dragonflies, which exhibit both exploration and exploitation behaviors. The algorithm begins with an initial population of dragonflies, each representing a potential solution to the optimization problem. These dragonflies move in the search space, exploring different solutions and exploiting promising areas to find the optimal solution. The population of dragonfly demonstrates five behaviors: separation, alignment, cohesion, foraging and evading enemies or predators, the following mathematical models with Eqs (9) -(13) describe these behaviors.

- The separation (S), is to avoid collision in the neighborhood:

$$S_i = -\sum_{j=1}^{N} X - Xj \tag{9}$$

  where X represents the position of the current individual, Xj is the position in the neighborhood individual j, and N is the total number of close individuals.
- Alignment principal (A), represents the identical speed of the members existing in the same group, and it is expressed as follows:

$$A_i = \left(\sum_{j=1}^{N} Vj/N\right) \quad (10)$$

Where Vj is the velocity of the j^{th} neighboring individual.

- The cohesion (C): is relied on the Attraction of individuals towards the center of the mass of the neighborhood, and it is described as follows:

$$C_i = \left(\sum_{j=1}^{N} Xj/N\right) - X \quad (11)$$

- The foraging behavior (F), the following equation determines the attraction to a food source, with X^+ representing the location of the food.

$$Fi = X^+ - X \quad (12)$$

- The behavior of eluding enemies (E) is described as follows:

$$Ei = X + X^- \quad (13)$$

X^- Represents the position of the enemy.

The localization of an artificial dragonfly is the sum of its positions X and it step ΔX, expressed as:

$$X_{t+1} = X_t + \Delta X_{t+1} \quad (14)$$

and it is shown in the following equation:

$$\Delta X_{t+1} = sS_i + aA_i + cC_i + fF_i + eE_i + \omega \Delta X_t \quad (15)$$

Where s is the separation weight and S_i is the separation for X_t noting a is the alignment weight, A_i is the alignment of i^{th} individual, c indicates the cohesion weight, C_i is the cohesion of the i^{th} individual, F_i is the food attraction of i^{th} individual and f is the food parameter, e is the enemy factor and F_i is the enemy's location of individual dragonfly, ω is the inertia weight, and t is the iteration counter (Mirjalili 2016).

The advantages of using the Dragonfly Algorithm include:

- Efficiency: The Dragonfly Algorithm efficiently explores the search space and converges towards optimal solutions by leveraging the collective behavior and movement patterns of dragonflies.
- Flexibility: The algorithm can be easily adapted and applied to various optimization problems due to its adaptive behavior and ability to search in the full range of wavelengths.
- Robustness: The Dragonfly Algorithm is less likely to get trapped in local optima due to its combination of random search and directed flight strategies, making it a robust optimization technique.
- Scalability: The algorithm can handle large-scale optimization problems with a large number of variables and constraints, making it suitable for complex biomedical engineering applications (Mirjalili 2016).

The comparison of DA-SVM performance with existing approaches such as PSO-SVM GA-SVM algorithms, the results revealed that the DA-SVM achieved competitive results of searching the optimal SVM parameters to improve the classification accuracy (Tharwat et al. 2017).

2.3. Description of the datasets collection and computational methodology

In this work, experimental data of the absorption and scattering coefficients of human blood were compiled averages from literature, in case tabulated data, and using the Get Data Graph Digitizer program (Bosschaart et al. 2014; Roggan et al. 1999) to build the models. These data were considered sufficient in building a reliable model because of repeatability of the results. Their blood samples were obtained from adults and healthy donors, fresh erythrocytes were centrifugated three times and washed in phosphate buffer solution (300mosmol/L, pH7.4). The hematocrit, which serves as an essential marker for the concentration of red blood cells, was precisely measured by employing dilution techniques with phosphate buffer to yield a diverse range of values up to 70%. They used Light microscopy to make sure that the samples were not hemolyzed or deformed, and the blood was mixed gently with a laminar flow, and the oxygen saturation was controlled by a blood oximeter and varied by a continuous flow of a mixture of O_2, N_2, and CO_2 through a modified membrane oxygenator, it is noteworthy that the experimental set up was carried out at constant temperature 20 C°. It is known that the optical properties of scattering and absorption media cannot be measured directly, they were determined by a double integrating sphere technique, evaluating the diffuse backscattering R_d, the total transmission T_t.

It should be mentioned that all computations for this paper were done on MATLAB 2019b, each of the two datasets was split into two parts; 80% of the randomly chosen data points were used to train the DA-SVR model, while the remaining 20% were used to validate the performance of the models developed. 800 out of the 1000 data points were used to train the machine and the remaining 200 data points were

used to test the performance of the models. We used the radial basis function (RBF), particularly the Gaussian kernel type, renowned for its higher performance (Wang et al. 2003). The procedures followed by the proposed DA-SVM are presented in Figure 1. Before starting iterations, initialization DA parameters, the position vector and step vector is required, and setting the maximum number of iteration (100 iterations), population size (30), upper and lower boundaries of particles (500 and 300, respectively). The data normalized as follows, $f' = \frac{f - \min}{\max - \min}$, where f is the original value, *min* and *max* are the lower and upper bounds of the feature value, respectively, and f' is the scaled value. The DA is determined when the maximum number of iterations are attained. this unique optimization strategy was able to effectively optimize the parameters of our model and achieve superior performance compared to other traditional optimization techniques. After the full search, the optimal parameters of DA-SVM models are shown in Table 1.

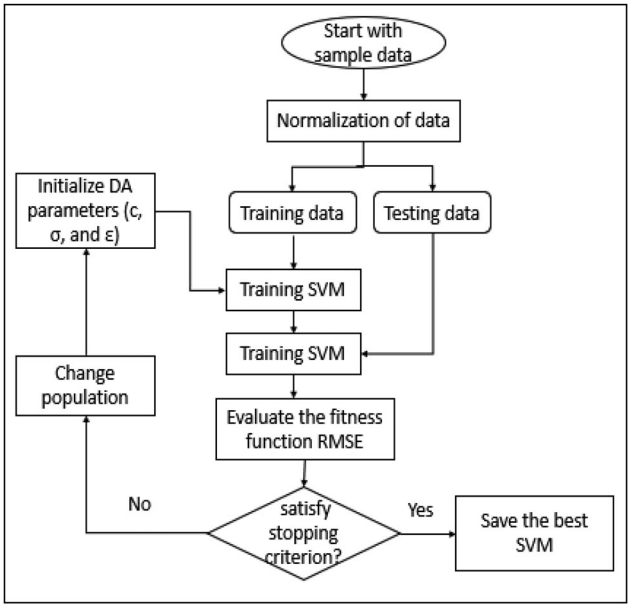


Figure 1. Flowchart of the DA-SVM process.

Table 1. Optimal DA-SVM parameters obtained.

Optimized SVR parameters	Absorption coefficient	Scattering coefficient
C	435	435
Kernel parameter	0.1	0.1046
Epsilon	0.6	0.6
Kernel function	Gaussian	Gaussian

3. Results

Proper use and interpretation of statistical parameters are crucial in research as they provide quantitative evidence to support research findings, facilitate comparisons, and help draw valid conclusions. The correlation coefficients (r) of the relationship between the features (independent variables) and the outputs (dependent variable) is an important statistical measure that can aid in feature selection, model performance assessment, and feature engineering, ultimately leading to better model interpretability, stability, and predictive accuracy (Table 2). As indicated, there is a negative relationship between the wavelength and the targets (μ_a and μ_s), and an extremely positive correlation coefficient (r) between the hematocrit and the target in both cases (μ_a and μ_s). The (r) values prove that there is a significant linear relationship between the input features variables and the targets (μ_a and μ_s) as long as *p*value are less than 0.05 and 0.01 which is a measure of the strength of evidence against the null hypothesis in a statistical hypothesis test. This indicates that these data could be used to build a reliable model for estimating the absorption and scattering coefficients of human blood. The mean or average, it is statistical parameter that provides a measure of central tendency, indicating the typical value in a dataset, is calculated by adding up all the values in a dataset and dividing by the number of values. The means and medians values of wavelength and hematocrit are very close (for both, absorption and scattering behaviors), it is an indication that the data are approximately symmetrically distributed, and the median value of absorption coefficient is closer to the lower quartile ($1.16\,mm^{-1}$), it suggested that the data may be skewed towards lower values, while the oxygen saturation value is closer to the upper quartile (89.62%), therefore, it may be skewed towards higher values.

The standard deviation measures the amount of variation in a dataset, there is larger standard

Table 2. Correlation coefficients and the respective *p*value of the relationship between the outputs and their descriptors.

	Absorption coefficient (μ_a)			Scattering coefficient (μ_s)		
Features	Wavelength vs μ_a	Hematocrit vs μ_a	Saturation of oxygen vs μ_a	Wavelength vs μ_s	Hematocrit vs μ_s	Saturation of oxygen vs μ_s
Correlation coefficient(r)	−0.4959	0.1872	0.0972	−0.7303	0.3623	0.26005
p-value	3.27×10^{-63}	2.44×10^{-09}	0.002	2.09×10^{-167}	2.15×10^{-32}	6.40×10^{-17}

Table 3. Statistical description of the absorption data adjusted in this work.

Variable	Unit	No of Data	Mean	Median	Minimum	Maximum	Standard deviation
Wavelength	nm	1000	999.030617	907.5	251	2220.8	524.300445
Hematocrit	%	1000	38.2888103	45	0.41352	56.7984	14.8005159
Oxygen saturation	%	1000	50.3865959	89.62765	0	100	49.318753
Absorption coefficient	mm^{-1}	1000	12.6135352	1.16	0.0105712	159.31	24.5790358

Table 4. Statistical description of the scattering data used in building the model.

Variable	Unit	No of Data	Mean	Median	Minimum	Maximum	Standard deviation
Wavelength	nm	1000	951.61995	792.5	272	1999.54	511.802475
Hematocrit	%	1000	38.7213047	45	0	69.7943	14.6956736
Oxygen saturation	%	1000	46.629	0	0	99	49.4413977
Scattering coefficient	mm^{-1}	1000	55.0827478	45.9794	0	128.16	30.2707498

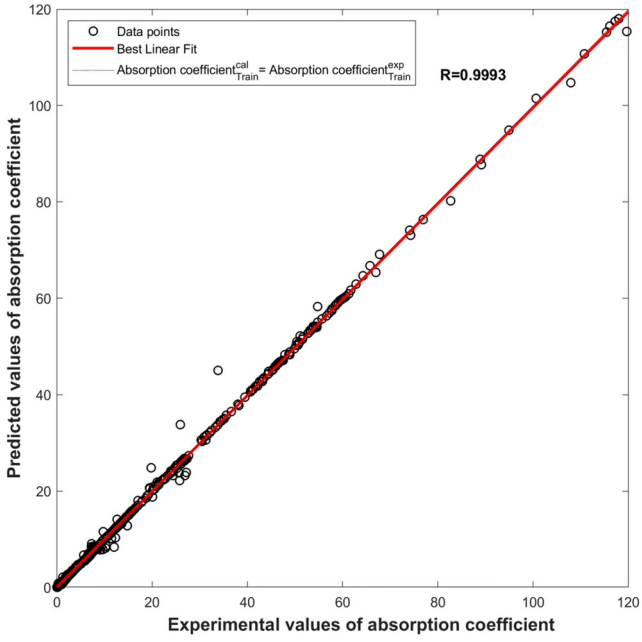


Figure 2. Scatter plots between experimental and DA-SVM predicted values for the training phase.

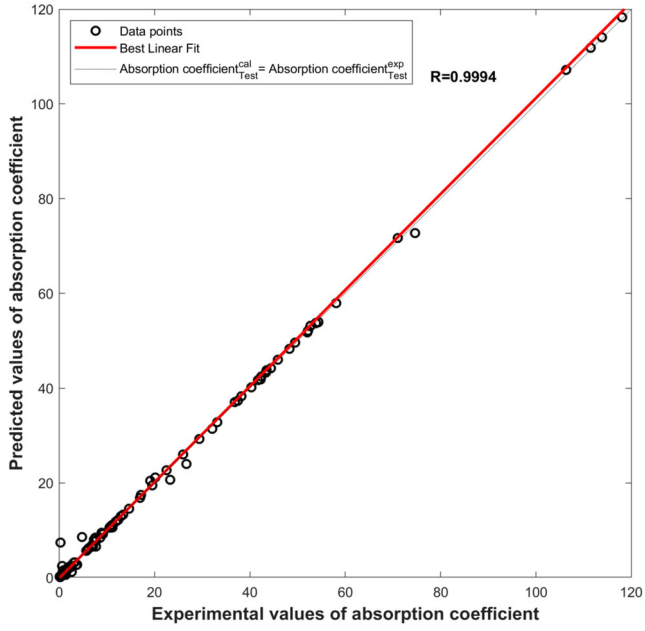


Figure 3. Scatter plots between experimental and DA-SVM predicted values for the testing phase.

deviation in the wavelength and hematocrit values, that showed greater variability in the data, while the oxygen saturation has a small standard deviation.

The descriptive statistics of these features are shown in Tables 3 and 4.

4. Discussion

To provide a comprehensive evaluation of the optimized DA-SVM models, we utilized the MATLAB function 'postreg' to plot the correlation between experimental and predicted values of absorption coefficient for both training and testing datasets. The resulting correlation coefficients were 0.9993 and 0.9994, respectively, as shown in Figures 2 and 3. These plots demonstrate the excellent performance of the models during both the training and testing phases. While the cross plots of measured and predicted values of scattering coefficient for the training and testing phases are shown in Figures 4 and 5. The plots showed correlation coefficients (R) of 0.9996 and 0.9957 for the training and testing datasets, respectively. Notably, the models displayed comparable prediction performance for the training phase, with correlation coefficients (R) values of 0.9993 and 0.9996. The correlation coefficients being close to unity indicate a strong agreement between the predicted and experimental results. This is further demonstrated by the good alignment of the predicted data set with all the experimental data on the first bisector regression, confirming the high level of accuracy achieved by the models.

The applicability domain (AD) of a prediction is also an important step in proving the reliability of the model, therefore the applicability domain is interpreted using William's plot (Figures 6 and 7). It is obvious that most samples in the dataset are within

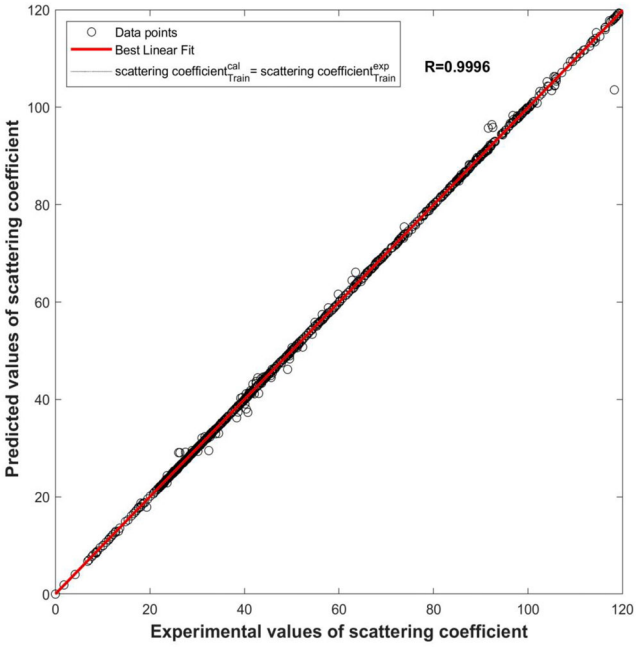


Figure 4. Correlation plot between the experimental and DA-SVM predicted values of training dataset.

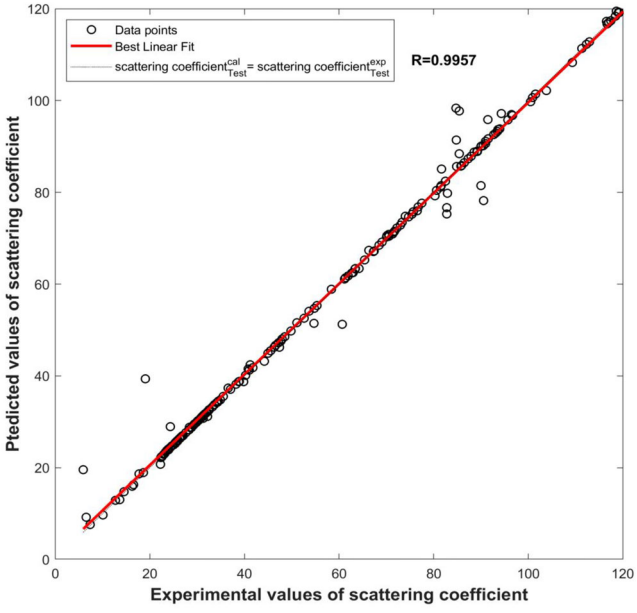


Figure 5. Correlation plot between the experimental and DA-SVM predicted values of testing dataset.

the applicability domain of the model with warning leverage limit is almost 0.013, and no more than 36 data points are outliers out of the 1000 overall experimental data for both of μ_a and μ_s.

In addition, assessing the model's predictive accuracy using statistical parameters such as the correlation coefficient (R), the root mean squared error ($RMSE$), and the mean absolute error (MAE). The mathematical formulas for these parameters are given below;

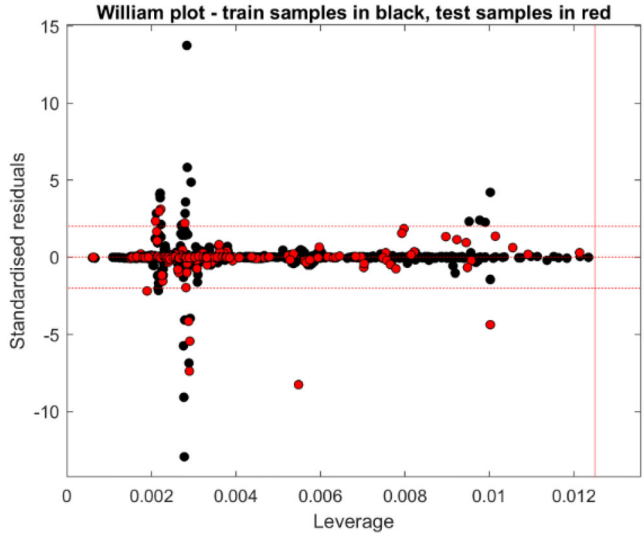


Figure 6. William's diagram for detection outliers for DA-SVM model (for μ_a).

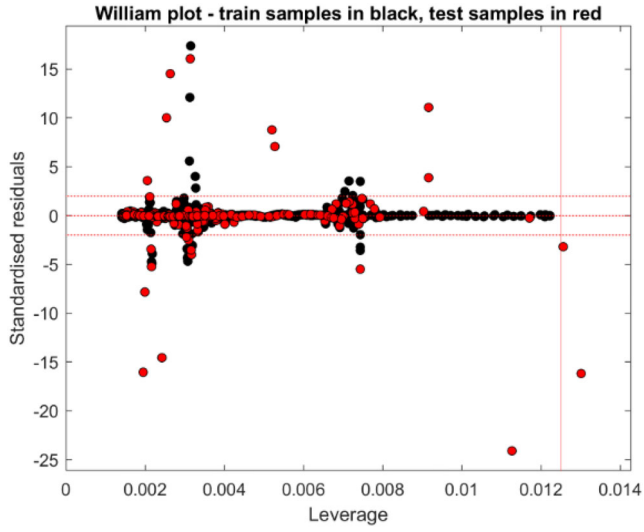


Figure 7. Williams plots to evaluate the applicability domain of the DA-SVM model (for μ_s).

$$R = \frac{\sum_i^N \left[\left(y_i^{exp} - \hat{y}^{exp}\right)\left(y_i^{cal} - \hat{y}^{cal}\right)\right]}{\sqrt{\sum_{i=1}^N \left(y_i^{exp} - \hat{y}^{exp}\right)^2}\sqrt{\sum_{i=1}^N \left(y_i^{cal} - \hat{y}^{cal}\right)^2}} \quad (16)$$

$$RMSE = \sqrt{\sum_{i=1}^N \left(y_i^{\exp} - y_i^{cal}\right)^2 / N} \quad (17)$$

$$MAE = (100/N) * \sum_{i=1}^N \left(\left|\left(\left(y_i^{\exp} - y_i^{cal}\right)/y_i^{\exp}\right)\right|\right) \quad (18)$$

Where N is the number of the dataset, y^{exp}, y^{cal} are the experimental and predicted values, respectively, and $\hat{y}^{\exp}$ is the mean of the experimental values, with

Table 5. Statistical results for the absorption and scattering coefficients of blood.

	Absorption coefficient		Scattering coefficient	
Statistics	Training	Testing	Training	Testing
Correlation Coefficient (*R*)	0.9993	0.9994	0.9996	0.9956
RMSE	0.8678	0.9792	0.8484	2.9193
MAE	0.2173	0.3509	0.2423	0.9692

$\hat{y}^{\exp} = \sum_{i=1}^{N} y_i^{\exp}/N$, the statistical results of the training and testing phases obtained during the development of the models (Table 5) revealed high correlation coefficients (close to 1). Additionally, the models demonstrated low root mean square error (*RMSE*) values of 0.8678 and 0.8484, as well as low mean absolute error (*MAE*) values of 0.2173 and 0.2423, for the training phases, respectively. These results showcase the robustness of the computational models and confirm the validity of the optimized DA-SVM models.

5. Conclusion

In this paper, we have successfully developed a new strategy to estimate the absorption (μ_a) and scattering properties of blood (μ_s) using computational intelligence. The proposed DA-SVR models used both physiological and biochemical factors such as the wavelength, the hematocrit percentage, and the oxygen saturation as descriptors in building the models The results of the Dragonfly Algorithm/Support Vector Regression models showed good correlative (*R*) with an accuracy of 0.9994 and 0.9956 for μ_a and μ_s, respectively, and low *RMSE* values were achieved. This study is pioneering the use of artificial intelligence to estimate the absorption and scattering properties of human blood, because of the excellent performance and high accuracy, we expect that this work will be helpful and opens up the horizons to the researchers to use computational intelligence approach to other optical properties of blood and even other human tissues.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The authors received no funds or grants for this work.

References

Alade IO, Bagudu A, Oyehan TA, Rahman MAA, Saleh TA, Olatunji SO. 2018. Estimating the refractive index of oxygenated and deoxygenated hemoglobin using genetic algorithm – support vector regression model. Comput Methods Programs Biomed. 163:135–142. doi:10.1016/j.cmpb.2018.05.029.

Alade IO, Abd Rahman MA, Saleh TA. 2019. Modeling and prediction of the specific heat capacity of Al2 O3/water nanofluids using hybrid genetic algorithm/support vector regression model. Nano-Structures Nano-Objects. 17:103–111. doi:10.1016/j.nanoso.2018.12.001.

Al-Jamimi HA, Al-Azani S, Saleh TA. 2018. Supervised machine learning techniques in the desulfurization of oil products for environmental protection: a review. Process Safety Environmental Protection. 120:57–71. doi:10.1016/j.psep.2018.08.021.

Al-Jamimi HA, Bagudu A, Saleh TA. 2019. An intelligent approach for the modeling and experimental optimization of molecular hydrodesulfurization over AlMoCoBi catalyst. J Molecular Liquids. 278:376–384. doi:10.1016/j.molliq.2018.12.144.

Alizamir A, Gholami A, Bahrami N, Ostadhassan M. 2022. Refractive Index of Hemoglobin Analysis: a Comparison of Alternating Conditional Expectations and Computational Intelligence Models. ACS Omega. 7(38):33769–33782. doi:10.1021/acsomega.2c00746.

Amari S, Wu S. 1999. Improving support vector machine classifiers by modifying kernel functions. Neural Netw. 12(6):783–789. doi:10.1016/s0893-6080(99)00032-5.

Benimam H, Moussa CS, Hentabli M, Hanini S, Laidi M. 2020. Dragonfly-Support Vector Machine for Regression Modeling of the Activity Coefficient at Infinite Dilution of Solutes in Imidazolium Ionic Liquids Using σ-Profile Descriptors. J Chem Eng Data. 65(6):3161–3172. doi:10.1021/acs.jced.0c00168.

Boser BE, Vapnik VN, Guyon IM. 1992. Training Algorithm Margin for Optimal Classifiers. Perception 144–152.

Bosschaart N, Edelman GJ, Aalders MCG, Van Leeuwen TG, Faber DJ. 2014. A literature review and novel theoretical approach on the optical properties of whole blood. Lasers Med Sci. 29(2):453–479. doi:10.1007/s10103-013-1446-7.

Byvatov E, Fechner U, Sadowski J, Schneider G. 2003. Comparison of Support Vector Machine and Artificial Neural Network Systems for Drug/Nondrug Classification. J Chem Inf Comput Sci. 43(6):1882–1889. doi:10.1021/ci0341161.

Chatterjee S, Biswas S, Majee A, Sen S, Oliva D, Sarkar R. 2022. Breast cancer detection from thermal images using a Grunwald-Letnikov-aided Dragonfly algorithm-based deep feature selection method. Comput Biol Med. 141:105027. doi:10.1016/j.compbiomed.2021.105027.

Drucker H, Surges CJC, Kaufman L, Smola A, Vapnik V. 1997. Support vector regression machines. Adv. Neural Inf. Process. Syst. 1:155–161.

Faber DJ, Aalders MCG, Mik EG, Hooper BA, Van Gemert MJC, Van Leeuwen TG. 2004. Oxygen saturation-dependent absorption and scattering of blood. Phys Rev Lett. 93(2):028102. doi:10.1103/PhysRevLett.93.028102.

Friebel M, Meinke M. 2005. Determination of the complex refractive index of highly concentrated hemoglobin solutions using transmittance and reflectance measurements. J Biomed Opt. 10(6):064019. doi:10.1117/1.2138027.

Friebel M, Roggan A, Müller G, Meinke M. 2006. 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. 11(3):34021. doi:10.1117/1.2203659.

Friebel M, Helfmann J, Netz U, Meinke M. 2009. Influence of oxygen saturation on the optical scattering properties of human red blood cells in the spectral range 250 to 2000 nm. J Biomed Opt. 14(3):034001. doi:10.1117/1.3127200.

Fu F, Lin J, Xu X. 2011. Effect of pH on optical properties during blood coagulation progress by optical coherence tomography. 4th International Conference on Biomed Eng Inf (BMEI) 1: p. 268–271.

Hammer M, Schweitzer D, Michel B, Thamm E, Kolb A. 1998. Single scattering by red blood cells. Appl Opt. 37(31):7410–7418. doi:10.1364/ao.37.007410.

Hariharan M, Sindhu R, Vijean V, Yazid H, Nadarajaw T, Yaacob S, Polat K. 2018. Improved binary dragonfly optimization algorithm and wavelet packet based non-linear features for infant cry classification. Comput Methods Programs Biomed. 155:39–51. doi:10.1016/j.cmpb.2017.11.021.

He Z, Wen X, Liu H, Du J. 2014. A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. J Hydrol. 509: 379–386. doi:10.1016/j.jhydrol.2013.11.054.

Hung FH, Chiu HW. 2017. Cancer subtype prediction from a pathway-level perspective by using a support vector machine based on integrated gene expression and protein network. Comput Methods Programs Biomed. 141:27–34. doi:10.1016/j.cmpb.2017.01.006.

Ibrir A, Kerchich Y, Hadidi N, Merabet H, Hentabli M. 2021. Prediction of the concentrations of PM1, PM2.5, PM4, and PM10 by using the hybrid dragonfly-SVM algorithm. Air Qual Atmos Health. 14(3):313–323. doi:10.1007/s11869-020-00936-1.

Jacques SL. 2013. Erratum: optical properties of biological tissues: a review. Phys Med Biol. 58(14):5007–5008. doi:10.1088/0031-9155/58/14/5007.

Jin YL, Chen JY, Xu L, Wang PN. 2006. Refractive index measurement for biomaterial samples by total internal reflection. Phys Med Biol. 51(20):N371–N379. doi:10.1088/0031-9155/51/20/N02.

Li LL, Zhao X, Tseng MLT, Tan RR. 2020. Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. J Clean Prod. 242:118447. doi:10.1016/j.jclepro.2019.118447.

Meinke M, Müller G, Helfmann J, Friebel M. 2007a. Optical properties of platelets and blood plasma and their influence on the optical behavior of whole blood in the visible to near infrared wavelength range. J Biomed Opt. 12(1):014024. doi:10.1117/1.2435177.

Meinke M, Müller G, Helfmann J, Friebel M. 2007b. Empirical model functions to calculate hematocrit-dependent optical properties of human blood. Appl Opt. 46(10):1742–1753. doi:10.1364/ao.46.001742.

Mirjalili S. 2016. Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. Neural Comput & Applic. 27(4):1053–1073. doi:10.1007/s00521-015-1920-1.

Mordon S, Rochon P, Dhelin G, Lesage JC. 2005. Dynamics of temperature dependent modifications of blood in the near-infrared. Lasers Surg Med. 37(4):301–307. doi:10.1002/lsm.20238.

Moussa H, Dahmoune F, Hentabli M, Remini H, Mouni L. 2022. Optimization of ultrasound-assisted extraction of phenolic-saponin content from Carthamus caeruleus L. rhizome and predictive model based on support vector regression optimized by dragonfly algorithm. Chemom Intell Lab Syst. 222:104493. doi:10.1016/j.chemolab.2022.104493.

Nilsson AMK, Lucassen GW, Verkruysse W, Andersson-Engels S, Van Gemert MJ. 1997. C. Changes in optical properties of human whole blood in vitro due to slow heating. Photochem Photobiol. 65(2):366–373. doi:10.1111/j.1751-1097.1997.tb08572.x.

Oduncuoğlu M. 2021. Refractive Index Formula of Blood as a Function of Temperature and Concentration. An Acad Bras Cienc. 93:1–8.

Oyehan TA, Alade IO, Bagudu A, O Sulaiman K, Olatunji SO, Saleh TA. 2018. Predicting of the refractive index of haemoglobin using the Hybrid GA-SVR approach. Comput Biol Med. 98:85–92. doi:10.1016/j.compbiomed.2018.04.024.

Prince S, Malarvizhi S. 2010. Estimation of optical properties of normal and diseased tissue based on diffuse reflectance spectral model. WCE 2010 - World Congr Eng 2010 1: 578–582.

Rakhshani MR, Mansouri-Birjandi MA. 2017. High sensitivity plasmonic refractive index sensing and its application for human blood group identification. Sensors Actuators, B Chem. 249:168–176. doi:10.1016/j.snb.2017.04.064.

Roggan A, Friebel M, Do Rschel K, Hahn A, Mu Ller G. 1999. Optical properties for circulating human blood in thewavelength range 400– 2500 nm. J Biomed Opt. 4(1): 36–46. doi:10.1117/1.429919.

Saleh TA. 2021. Protocols for synthesis of nanomaterials, polymers, and green materials as adsorbents for water treatment technologies. Environmental Technology Innovation. 24:101821. doi:10.1016/j.eti.2021.101821.

Sardar DK, Levy LB. 1998. Optical properties of whole blood. Lasers Med Sci. 13(2):106–111. doi:10.1007/s101030050062.

Sardar DK, Swanland GY, Yow RM, Thomas RJ, Tsin ATC. 2007. Optical properties of ocular tissues in the near infrared region. Lasers Med Sci. 22(1):46–52. doi:10.1007/s10103-006-0421-y.

Smola AJ, Schölkopf B, Müller KR. 1998. The connection between regularization operators and support vector kernels. Neural Netw. 11(4):637–649. doi:10.1016/s0893-6080(98)00032-x.

Steinke JM, Shepherd AP. 1992. Effects of temperature on optical absorbance spectra of oxy-, carboxy-, and deoxy-hemoglobin. Clin Chem. 38(7):1360–1364. doi:10.1093/clinchem/38.7.1360.

Tharwat A, Gabel T, Hassanien AE. 2018. Parameter optimization of support vector machine using dragonfly algorithm. In: Hassanien A, Shaalan K, Gaber T, Tolba M, editors. Proceedings of the International Conference on Advanced Intelligent Systems and Informatics. p. 309–319. Cham: Springer International Publishing. doi:10.1007/978-3-319-64861-3_29

Vladimir NVapnik. 1995. The nature of staristical learning theory, Discov Data Min Knowl 1–47.

Wang J, Shi Z, Fan J, Ge Y, Yin J, Hu G. 2012. Self-assembly of graphene into three-dimensional structures promoted by natural phenolic acids. J Mater Chem. 22(42): 22459–22466. doi:10.1039/c2jm35024f.

Wang W, Xu Z, Lu W, Zhang X. 2003. Determination of the spread parameter in the Gaussian kernel for classification and regression. Neurocomputing. 55(3-4):643–663. doi:10.1016/S0925-2312(02)00632-X.

Xie T, Yao J, Zhou Z. 2019. DA-based parameter optimization of combined kernel support vector machine for cancer diagnosis. Processes. 7(5):263. doi:10.3390/pr7050263.

Yahya M, Saghir MZ. 2016. Empirical modelling to predict the refractive index of human blood. Phys Med Biol. 61(4):1405–1415. doi:10.1088/0031-9155/61/4/1405.

Yaroslavsky AN, Yaroslavsky IV, Goldbach, T, Schwarzmaier H J. 1996. Optical properties of blood in the near-infrared spectral range. Proceedings Volume 2678, Optical Diagnostics of Living Cells and Biofluids. San Jose, CA: Photonics West '96. doi:10.1117/12.239516.

Zhernovaya O, Bashkatov AN, Genina EA, Tuchin V, Meglinski IV, Churmakov DY, Ritchie LJ. 2007. Investigation of glucose-hemoglobin interaction by optical coherence tomograph. In. Saratov Fall Meeting: optical Technologies in Biophysics and Medicine. VIII. 6535:423–429.

Zhernovaya O, Sydoruk O, Tuchin V, Douplik A. 2011. The refractive index of human hemoglobin in the visible range. Phys Med Biol. 56(13):4013–4021. doi:10.1088/0031-9155/56/13/017.