ABSTRACT

The aim of this study is to predict anemia from a population through biomedical variables by using the optimum linear model. A linear medical model based on biomedical variables is produced and an effective technique is used in investigating the optimum parameters of the model. To achieve this, the particle swarm optimization (PSO) algorithm have effectively been applied in predicting the parameters of the model through the biomedical variables. The study is conducted in terms of data set consisting of 539 subjects provided from blood laboratories. Optimum values of the parameters produced from the PSO algorithm are used here to obtain the more realistic model. The model based on the variables and outcomes is expected to serve as a good indicator of disease diagnosis for health providers and planning treatment schedules for their patients. Thus, the article is believed to be beneficial especially for who are interested in biomedical models arising in various fields of medical science, especially anemia.

Keywrods: Anemia, medical model, linear model, particle swarm optimization.

1. INTRODUCTION

Anemia is defined clinically as blood hemoglobin value that is below the appropriate reference range for the individual. This decrease in the level of hemoglobin leads to the lack of access to the tissue a sufficient amount of oxygen and therefore appear in the symptoms of a headache, fatigue, inability to focus, attention, weakness, exhaustion, chest pain, cold hands, and feet. Anemia is one of the most common causes of blood diseases worldwide. There are many different types of anemia. Depending on the type, the symptoms of anemia can range from short episodes to chronic conditions. Each type of anemia produces a different case, ranging from moderate to severe and each has its own causes. Anemia can be either temporary or long-term [1-5].

The progress of medical models considered to produce medical outputs is important tools to deal with the behavior of a medical problem. They depend on the quality of any particular objective achieved on the state of knowledge about the system and how well successful modeling. As indicated in the literature [5-11], mathematical medical modelling has been realized to be a fundamentally important tool for the analysis of pathological characteristics. Response to a
medical model to limits of performance is of major interest and thus the current medical model describes the relationship between the biomedical variables and the diseases. The observational data may be modelled by a function linearly. Here the parameters for each of the variables in the linear medical model are estimated that to be the optimal model for more accurate prediction of anemia through the biomedical information.

Many models have been produced in dealing with various medical problems in the literature such as congenital heart disease [12], diabetic nephropathy [13], osteoporosis [14], and cancers [15, 16]. A frequently encountered medical problem is that of having a set of data, which one wishes to describe it by a mathematical model and determine a set of parameters that characterize the model. In this study, the major emphasis will be the fitting parameters of the model assumed to have some particular medical or mathematical significance through estimating best values in the set of the parameters. Therefore, the main aim here is to develop a medical model to study the effect of the blood variables, sex, and age on the pathologies through a large group of the variables because there has been an increase in the incidence of anemia among different segments of society.

Some other estimation methods [17-20] to analyze disease problems in addition to anemia. Heuristic algorithms can be effectively used to find the optimal parameters for the linear model in plenty of medical studies. Therefore, the PSO is one of the most efficient optimization algorithms that are used for a wide range of complex optimization problems. In computational science, the PSO is a computational method that works to improve the problem by repeatedly trying to improve the candidate solution. Therefore, these candidate solutions are created by the method repeatedly for improving the possibility of being the actual solution.

The PSO proposed by Kennedy and Earhart [21] has been used to solve various optimization problems in the literature [22-33], to estimate the parameters of models and implemented different strategies of mathematical methods to predict and to optimize problems.

Despite the recognized advantages of conventional methods, most of them suffer from various disadvantages such as high cost, difficulty in use, and time-consuming. In this case, optimization can be recalled as a very good alternative to the corresponding methods. In the past several years, the PSO has been successfully applied to areas to simplify optimization problems that had previously experienced serious difficulties. It is demonstrated that the PSO gets better results in a faster, cheaper way and the simplicity of the implementation, which are the most attractive features of this algorithm. Another reason that makes the PSO attractive is that reliable, robust, and considered as an effective meta-heuristic optimization algorithm. The PSO inspired by the behaviour of social models for flocking birds or fish education are based on individual improvement and social collaboration [34-38]. In this study, the PSO approach has been proposed to estimate the best parameter values of the linear medical model. This algorithm is common in the academic community as a typical tool because of its ability to optimize complex search spaces. Thus, the above advantages of the PSO sent us to use in dealing with the current medical problem. It should be borne in mind that fewer blood variables may cause the problem not to be effectively represented.

This paper is structured as follows. The next section discusses the study samples of the medical dataset, explain the linear model procedure, the PSO algorithm, and how to test the model and estimate parameters of the linear medical model. Section 3 presents the results and discussion. Finally, conclusions and recommendation for future work have been detailed.

2. MATERIALS AND METHODS

2.1. Study samples of the medical dataset

The data used here were collected from observations of anemia and included (539 subjects, 211 healthy subjects, 328 sick subjects) provided from blood laboratories in Iraq and we have
taken observations of the ages of individuals between (6-56) years. Here, we have some blood
diseases are Iron deficiency anemia (1), Deficiency Vitamin B12 (2), Thalassemia (3), Sickle cell
(4) and Spherocytosis (5). For each disease, we have samples for the individuals and for each
individual readings of the blood variables are Hemoglobin (HB), Red Blood Cell (RBC), Mean
Corpuscular Hemoglobin (MCH), White Blood Cell (WBC), Mean Corpuscular Volume (MCV),
Haematocrit (HCT), Mean Corpuscular Hemoglobin Concentration (MCHC), Platelets (PLT), and
sex (male (1) and female (2)), and age. The number of variables studied for the model is
consisting of ten independent variables and a dependent variable. The dependent variable consists
of six different types of output (healthy subject: 0 and blood diseases: 1-5).

2.2. Modelling

A linear model is an engine behind a multitude of data applications used for many forms of
prediction. Therefore, processes are governed by linear models in various fields of science such as
the estimation of the parameters of a linear medical model for predicting anemia.

A linear medical model describes a linear relationship between the dependent and
independent variables. The derived model is as follows:

$$\mathbf{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \epsilon. \quad (1)$$

The linear model with \( k \) predictor variables and the observations recorded for each of these \( n \)
levels can be expressed in the following style:

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_k x_{1k} + \epsilon_1$$

$$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_k x_{2k} + \epsilon_2$$

$$\vdots$$

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i$$

$$\vdots$$

$$y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} + \epsilon_n. \quad (2)$$

Here \( y_1, y_2, \ldots, y_n \) and \( x_1, x_2, \ldots, x_k \) stand for the dependent and independent observations,
respectively.

System (2) can be reexpressed in a more compact way:

$$\mathbf{y} = \beta \mathbf{X} + \epsilon, \quad (3)$$

with

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \text{ and } \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \quad (4)$$

where \( \mathbf{y}, \mathbf{X}, \beta \) and \( \epsilon \) indicate to the observations, the parameters of the model and the
unobserved random variable that adds noise to the linear relationship, respectively.

To obtain the linear model, \( \beta \) should be known. \( \hat{\beta} \) is estimated by minimizing the sum of the
squared error function \( SSE(\beta) \) under the consideration of the PSO. Knowing the estimates \( \hat{\beta} \), the
linear model can now be expressed as [39,40]

$$\hat{\mathbf{y}} = \hat{\beta} \mathbf{X} \quad (5)$$

where \( \hat{\mathbf{y}} \) is the estimated value for \( \mathbf{y} \).
2.3. Particle Swarm Optimization

The PSO is a population-based stochastic approach, invented by Eberhart and Kennedy [21], for solving continuous and discrete problems. They inspired from social behavior of bird flocking or fish schooling, these animals have a major role in the development of the algorithm.

The method optimizes a problem by trying to improve a solution. Each particle traces its coordinates in the area of the problem that relates to the best solutions carried out so far. This value is called pbest. Another "best" value that is tracked by the PSO is the best value, obtained so far by any particle in the neighbors of the particle. This location is called lbest. When the particle considers the whole population as its topological neighbors, the best value is a global best and is called gbest. The PSO idea consists of, at each time step, changing the velocity of each particle towards the pbest and lbest locations.

In the PSO, simple software agents, called particles, move in the search space for improvement. These randomly selected particles search solution space using the information of their neighborhood, personal information, and randomness. The position of a particle represents a candidate solution to the existing improvement problem. All particles look for better sites in the search space by changing their velocity at the end of each iteration. Because of each iteration, the position and velocity vectors are expressed as follows:

\[
V_{i}^{t+1} = \omega V_{i}^{t} + c_{1}r_{1}(P_{best} - X_{i}^{t}) + c_{2}r_{2}(G_{best} - X_{i}^{t}) \tag{6}
\]

\[
X_{i}^{t+1} = X_{i}^{t} + V_{i}^{t+1} \tag{7}
\]

where \( t, \omega, c_{1}, c_{2}, r_{1}, r_{2}, V_{i}^{t}, X_{i}^{t}, P_{best} \) and \( G_{best} \) indicate iteration number, weight parameter, acceleration coefficients (cognitive parameter, social parameter), random numbers uniformly distributed between 0 and 1, velocity of individual \( i \) at iteration \( t \), position of individual \( i \) at iteration \( t \), the best local value of each particle, the best value of swarm, respectively [38,41,42].

2.4. Test for the model

The coefficient of the determination, usually referred to as \( R^2 \), is a measure explaining the change in the relationship between all blood variables, sex, and age and the anemia types.

Here, we present some initial considerations. Consider the variance of the observations \( y \) by analyzing the total sum of squares, denoted by \( SST \) and the sum of squared errors, denoted by \( SSE \). That is,

\[
SST = \sum_{j=1}^{n}(y_j - \bar{y})^2 \tag{8}
\]

and

\[
SSE = \sum_{j=1}^{n}(y_j - \hat{y}_j)^2 = \sum_{j=1}^{n} e_j^2 \tag{9}
\]

Now, the coefficient of the determination is defined by

\[
R^2 = \frac{SST - SSE}{SST} \tag{10}
\]

If the percentage explained by the coefficient of the determination is small, compatibility may not be very appropriate.

A terminological difference arises in the expression root mean squared error (RMSE). It is the square root of the average squared differences between the prediction and actual observations. The RMSE indicate the concentration of data around the model. In other words, it tells us how the data is centered around the most appropriate line [39,40,43]. It is very common to use the RMSE in the predictions. Then it is given by

\[
RMSE = \sqrt{MSE} \tag{11}
\]

Thus, it is given by
\[ \text{MSE} = \frac{1}{n} \sum_{j=1}^{n} e_j^2. \] (12)

2.5. Estimation of the parameters of the linear medical model

The currently linear medical model is a linear equation for our data. The model is as follows:
\[ y = \beta_0 + \beta_1 HB + \beta_2 RBC + \beta_3 MCH + \beta_4 WBC + \beta_5 MCV + \beta_6 HCT + \beta_7 MCHC + \beta_8 PLT + \beta_9 Sex + \beta_{10} Age \] (13)

where \( y \) is the type of anemia and \( \beta_i \), \( 0 \leq i \leq 10 \), are the parameters to be determined.

Here HB, RBC, MCH, WBC, MCV, HCT, MCHC, PLT stand for Hemoglobin, Red Blood Cell, Mean Corpuscular Hemoglobin, White Blood Cell, Mean Corpuscular Volume, Haematocrit, Mean Corpuscular Hemoglobin Concentration, Platelets, respectively.

As previously mentioned, the model can be represented in a more compact form as follows:
\[ \hat{y} = \hat{\beta} x \] (14)

where
\[ \hat{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{539} \end{bmatrix}, \quad x = \begin{bmatrix} 1 & HB_{11} & RBC_{12} & \cdots & Age_{110} \\ 1 & HB_{21} & RBC_{22} & \cdots & Age_{210} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & HB_{539,1} & RBC_{539,2} & \cdots & Age_{539,10} \end{bmatrix}, \quad \hat{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{10} \end{bmatrix}. \] (15)

Here \( \hat{y}, x \) and \( \hat{\beta} \) represent the estimates for output (anemia), the independent observations matrix, and estimated parameters, respectively.

This study aims at estimating the parameters \( \beta \) by minimizing the sum of the squared error function \( \text{SSE}(\beta) \) under the consideration of the PSO.

Hence, the fitness function in the PSO search engine is selected as the \( \text{SSE}(\beta) \), specifically:
\[ \text{SSE}(\beta) = \sum_{i=1}^{n} (y_i - f(x_i, \beta))^2. \] (16)

For the linear model in equation (13),
\[ \text{SSE}(\beta) = \sum_{i=1}^{539} [y_i - (\beta_0 + \beta_1 HB + \beta_2 RBC + \beta_3 MCH + \beta_4 WBC + \beta_5 MCV + \beta_6 HCT + \beta_7 MCHC + \beta_8 PLT + \beta_9 Sex + \beta_{10} Age)]^2. \] (17)

Here \( y_i \) are the dependent observations, \( \beta_i, 0 \leq i \leq 10 \), are the parameters to be determined.

In this article, the PSO is effectively used to estimate the parameters of the linear medical model in deriving an accurate model by finding a rapid convergence of the minimum value of the sum of the squared error in fewer iterations provides accurate estimates for parameter estimation of the linear medical model (see Tables 1-5).

The settings for the main parameters of the PSO method (\( \omega, c_1, c_2 \), and the size of the swarm) determine how to optimize the search space. Usually decreases the parameter \( \omega \) from around 0.9 to around 0.4 during the computation, the appropriate value for the parameter \( \omega \) provides a balance between the global and local exploration capacity of the swarm and thus a better solution [22,35-37]. If the parameter \( \omega \) is much less than one, only a small momentum of the previous time step is preserved, thus rapid changes in the direction are possible with this setting. High settings near 1 facilitate global searching. The usual choices for acceleration coefficients are \( c_1 \) and \( c_2 \), usually, \( c_1 \) is equal to \( c_2 \) and ranges between 0 and 4. The size of swarm plays a very important role in the PSO, as is the durability and complexity of the algorithm. By inspiring from the literature [22,36,37], we have produced our PSO algorithm as given in Figure 1.
Figure 1. The PSO algorithm
3. RESULTS AND DISCUSSION

The current study focuses to obtain the best estimate of the parameters through the PSO for the currently derived linear model to detect the link between the biomedical variables and anemia.

As opposed to the PSO approach, classical methods in dealing with linear models have some disadvantages as seen in the previous works [29-33], where they require many mathematical operations; like the Jacobean matrix, and matrix operations.

The researchers estimated parameters of a great number of models by using the PSO in the literature [22-27]. They discussed different problems/models by using their own approaches. We have here studied a linear model for a great number of biomedical data of anemia through the PSO to estimate the parameters for the model and investigating the relationship between many blood variables and the anemia types as opposed to researchers in the literature [44-47], they used a very limited number of blood variables or a few the anemia types.

Here, we have estimated the parameters of the linear model through the PSO algorithm (see Tables 1-4), and the produced results for various versions of the model by the minimum error (see Table 5). In the estimation, when the number of iterations is increasing, the error is decreasing as seen in Figures 2-5. Notice that the iteration reaches its optimum level at 4500.

<table>
<thead>
<tr>
<th>Biomedical Variables</th>
<th>Parameters $\beta_i$, $0 \leq i \leq 10$</th>
<th>SST</th>
<th>SSE($\beta$)</th>
<th>RMSE</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-3.167</td>
<td>1157.243</td>
<td>1817.378</td>
<td>1.836</td>
<td>0</td>
</tr>
<tr>
<td>HB</td>
<td>-0.726</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBC</td>
<td>-0.634</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCH</td>
<td>0.901</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WBC</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCV</td>
<td>0.125</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCT</td>
<td>0.062</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCHC</td>
<td>-1.408</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLT</td>
<td>0.257</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0.465</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1.** Parameter estimation by the PSO algorithm when the iteration is 500

![Figure 2. Sum of squared errors of the PSO algorithm when the iteration is 500](image-url)
Table 2. Parameter estimation by the PSO algorithm when the iteration is 1000

<table>
<thead>
<tr>
<th>Biomedical Variables</th>
<th>Parameters $\beta_i, 0 \leq i \leq 10$</th>
<th>$SST$</th>
<th>$SSE(\beta)$</th>
<th>RMSE</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>7.799</td>
<td>1157.243</td>
<td>1080.449</td>
<td>1.416</td>
<td>0.066</td>
</tr>
<tr>
<td>HB</td>
<td>1.399</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBC</td>
<td>1.141</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCH</td>
<td>-0.275</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WBC</td>
<td>-0.018</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCV</td>
<td>0.114</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCT</td>
<td>-0.743</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCHC</td>
<td>-0.149</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLT</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0.596</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.023</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. Sum of squared errors of the PSO algorithm when the iteration is 1000

Table 3. Parameter estimation by the PSO algorithm when the iteration is 2000

<table>
<thead>
<tr>
<th>Biomedical Variables</th>
<th>Parameters $\beta_i, 0 \leq i \leq 10$</th>
<th>$SST$</th>
<th>$SSE(\beta)$</th>
<th>RMSE</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>5.603</td>
<td>1157.243</td>
<td>405.983</td>
<td>0.868</td>
<td>0.649</td>
</tr>
<tr>
<td>HB</td>
<td>0.252</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBC</td>
<td>0.270</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCH</td>
<td>0.146</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WBC</td>
<td>0.0002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCV</td>
<td>0.053</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCT</td>
<td>-0.304</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCHC</td>
<td>-0.038</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLT</td>
<td>0.0008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>-0.208</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.013</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4. Sum of squared errors of the PSO algorithm when the iteration is 2000

Table 4. Parameter estimation by the PSO algorithm when the iteration is 4500

<table>
<thead>
<tr>
<th>Biomedical Variables</th>
<th>Parameters $\beta_i, 0 \leq i \leq 10$</th>
<th>SST</th>
<th>$SSE(\beta)$</th>
<th>RMSE</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>6.345</td>
<td>1157.243</td>
<td>347.989</td>
<td>0.803</td>
<td>0.699</td>
</tr>
<tr>
<td>HB</td>
<td>-0.201</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBC</td>
<td>-0.461</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCH</td>
<td>-0.033</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WBC</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCV</td>
<td>0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCT</td>
<td>-0.022</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCHC</td>
<td>0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLT</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>-0.306</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5. Sum of squared errors of the PSO algorithm when the iteration is 4500
Table 5. Parameter estimation of the various forms by the PSO algorithm when the iteration is 4500

<table>
<thead>
<tr>
<th>Models</th>
<th>SSE</th>
<th>RMSE</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 for (HB, sex and age)</td>
<td>500.117</td>
<td>0.963</td>
<td>0.568</td>
</tr>
<tr>
<td>Model 2 for (RBC, sex and age)</td>
<td>956.017</td>
<td>1.332</td>
<td>0.174</td>
</tr>
<tr>
<td>Model 3 for (MCH, sex and age)</td>
<td>862.084</td>
<td>1.265</td>
<td>0.255</td>
</tr>
<tr>
<td>Model 4 for (WBC, sex and age)</td>
<td>891.756</td>
<td>1.286</td>
<td>0.229</td>
</tr>
<tr>
<td>Model 5 for (MCV, sex and age)</td>
<td>937.336</td>
<td>1.319</td>
<td>0.190</td>
</tr>
<tr>
<td>Model 6 for (HCT, sex and age)</td>
<td>406.077</td>
<td>0.868</td>
<td>0.648</td>
</tr>
<tr>
<td>Model 7 for (MCHC, sex and age)</td>
<td>876.008</td>
<td>1.275</td>
<td>0.243</td>
</tr>
<tr>
<td>Model 8 for (PLT, sex and age)</td>
<td>843.894</td>
<td>1.251</td>
<td>0.271</td>
</tr>
<tr>
<td>Model 9 for (HB, MCH, sex and age)</td>
<td>496.170</td>
<td>0.959</td>
<td>0.571</td>
</tr>
<tr>
<td>Model 10 for (RBC, WBC, sex and age)</td>
<td>885.520</td>
<td>1.282</td>
<td>0.235</td>
</tr>
<tr>
<td>Model 11 for (MCV, PLT, sex and age)</td>
<td>829.614</td>
<td>1.241</td>
<td>0.283</td>
</tr>
<tr>
<td>Model 12 for (MCHC, HCT, sex and age)</td>
<td>389.654</td>
<td>0.850</td>
<td>0.663</td>
</tr>
<tr>
<td>Model 13 for (HB, WBC, HCT, sex and age)</td>
<td>384.303</td>
<td>0.844</td>
<td>0.667</td>
</tr>
<tr>
<td>Model 14 for (MCV, MCHC, RBC, sex and age)</td>
<td>844.280</td>
<td>1.252</td>
<td>0.270</td>
</tr>
<tr>
<td>Model 15 for (HB, RBC, MCH, WBC, sex and age)</td>
<td>353.664</td>
<td>0.810</td>
<td>0.690</td>
</tr>
<tr>
<td>Model 16 for (MCV, HCT, MCHC, PLT, sex and age)</td>
<td>378.580</td>
<td>0.838</td>
<td>0.670</td>
</tr>
</tbody>
</table>

In this study, the size of the swarm is taken to be according to the structure of the linear medical model, the number of estimated parameters, and searching space between (-10 and 10). The acceleration coefficients; cognitive parameter $c_1$ and social parameter $c_2$ are selected as 1 and 3, respectively. The algorithm is set to stop after different iterations and different independent experiments to check the durability of the estimation strategy.

Estimating the parameters of the medical model is a difficult task for classical methods of optimization. The starting values for the parameters are randomly selected from the search area. The $\beta$ values refer to the estimated parameter values for the real parameters obtained by the PSO. After different independent attempts have been made and different iterations 500, 1000, 2000 and 4500 have been taken to obtain the best parameters, and then we have obtained the best estimated parameters with iterations of 4500 (see Tables 1-4 and Figures 2-5).

Since the PSO algorithm is random inherently, convergence behavior and final estimated values can be of attention. For the medical model, the behavior of the error function is interpreted through the PSO approach, which consists of the values evaluated during the process of minimization (see Figures 2-5).

The parameter value is suitable for the model, when $SSE = 347.989$, $RMSE = 0.803$, and $R^2 = 0.699$ by the PSO. This is important because the SSE measures how well the data fit the model and means a better fit the model with the data and small values of the $RMSE$ indicate the concentration of data around the model line. The medical model of interest has been seen to be effective significantly, on the prediction of the anemia types, which explain 69.90% of the change in the relation of the model between the observational variables and the anemia types.

The results obtained from the $SSE, RMSE, and R^2$ by using the PSO at the iteration of 4500, that the models produced in terms of a great number of blood variables a better relationship appear than the models produced in terms of fewer number of blood variables for predicting the anemia types (see Tables 4,5).
4. CONCLUSIONS AND RECOMMENDATION

This study has discovered the anemia types through biomedical information under the consideration of eight different blood variables, sex, and age of individuals. Therefore, it has developed an alternative for estimating the parameter approach that depends on the PSO algorithm in a linear medical model. As opposed to classical methods, it has been seen that the PSO approach is more advantageous, it requires less mathematical operations to estimate medical model parameters. It can be concluded that the PSO algorithm has been considered as an effective and very appropriate estimating method for the current and similar to current medical models. The parameter values produced are seen to be the most up-to-date and maybe the best. Thus, the PSO algorithm shows the tendency of rapid convergence for the model with the knowledge that the number of parameters is eleven. For further study, to estimate the parameters of the medical model, various computational methods can be analyzed.

REFERENCES


