Optimizing the Design of a TIR Lens Using SVR, VIKOR, and the Artificial Bee Colony Algorithm

Chih-Ming Hsu

Abstract—Multi-response parameter design problems have become increasingly important and have received considerable attention from both researchers and practitioners since there are usually several quality characteristics that must be optimized simultaneously in most modern products/processes. This study applies support vector regression (SVR), VlseKriterijumska Optimizacija I Kompromisno Resenje (VIKOR), and the artificial bee colony (ABC) algorithm to resolve these common and complicated parameter design problems. The feasibility and effectiveness of the proposed approach are demonstrated via a case study in which the design of a total internal reflection (TIR) lens is optimized while fabricating an MR16 light-emitting diode (LED) lamp. Experimental results indicate that the proposed solution procedure can provide highly robust design parameter settings for TIR lenses that can be directly applied in real manufacturing processes. The comparison revealed that the ABC algorithm can search the solution spaces in continuous domains modeled via SVR instead of in the limited discrete experiment levels thus finding a more robust design.

Index Terms—total internal reflection lens, support vector regression, VIKOR, artificial bee colony, multi-response parameter design

I. INTRODUCTION

In many real world applications, several output responses of a product/system must be optimized simultaneously through determining the optimal settings of input variables (control factors), called multi-response parameter design problems. The Taguchi method is a well-known traditional approach for addressing such problems; however, some subjective trade-offs must be made while selecting the optimal setting of each control factor in order to simultaneously consider all responses. Therefore, many approaches which combine techniques in various fields have suggested in the literature to tackle parameter design problems with multiple responses. For example, Kim and Lin [1] presented an approach that aims to maximize the overall minimal value of satisfaction with respect to all responses in order to address the multi-response parameter design problem by using response surface methodology (RSM) and exponential desirability functions. Lu and Antony [2] proposed the application of a fuzzy-rule based inference system and the Taguchi method to tackle multi-response optimization problems. Tong et al. [3] applied principal component analysis (PCA) and the technique for order preference by similarity to ideal solution (TOPSIS) to optimize multiple responses simultaneously. Kovach and Cho [4] developed a multidisciplinary-multiresponse robust design (MMRD) optimization approach for resolving parameter design problems with multiple responses using the combined array design and the nonlinear goal programming technique. Routara et al. [5] proposed an approach that applies weighted principal component analysis (WPCA), combined quality loss (CQL), and the Taguchi method to tackle multi-response optimization problems. Sibaliuja et al. [6] proposed an integrated approach based on Taguchi method, principal component analysis (PCA), grey relational analysis (GRA), neural networks (NNs), and genetic algorithms (GAs) to optimize a multi-response process. Hsu [7] developed a systematic procedure based on genetic programming (GP) and ant colony optimization (ACO) for optimizing the settings of control factors in multi-response parameter design problems. Al-Refaie [8] proposed a procedure that uses two techniques of data envelopment analysis (DEA) to improve the performance of a product/process with multiple responses. Salmasnia et al. [9] presented a three-phased approach that uses principal component analysis (PCA), adaptive-network-based fuzzy inference systems (ANFIS), desirability function and genetic algorithms (GAs) to simultaneously optimize multiple correlated responses in which the relationships between responses and design variables are highly nonlinear. He et al. [10] considered the uncertainty associated with the fitted response surface model by taking account of all values in the confidence interval rather than a single predicted value for each response to simultaneously optimize multiple responses. Bera and Mukherjee [11] proposed an adaptive penalty function-based “maximin” desirability index for multiple-response optimization (MRO) problems with close engineering tolerances of quality characteristics. In addition, a near-optimal solution for the single objective, i.e., desirability index, problem is determined via continuous ant colony optimization, ant colony optimization in real space, and global best particle swarm optimization.

In this study, a general procedure for resolving common, complicated multi-response parameter design problems based on support vector regression (SVR), VlseKriterijumska Optimizacija I Kompromisno Resenje (VIKOR), and the artificial bee colony (ABC) algorithm is proposed.
II. RESEARCH METHODOLOGIES

A. Support Vector Regression

The support vector regression (SVR) [12, 13] is the application of the support vector machine (SVM) to cases of function approximation or regression. Given a training data set \( \{x_i, y_i\}_{i=1}^n \), where the input variable \( x_i \in \mathbb{R}^n \) is an \( n \)-dimensional vector and the output variable \( d_i \in \mathbb{R} \) is a real value, we want to construct an appropriate model to describe the functional dependence of \( d \) on \( X \). SVR uses a map \( \Phi \) to transform a non-linear regression problem into a linear regression problem in a high dimensional feature space, and approximates a function of the form

\[
f(X, W) = \sum w_i \phi_i(X) + w_0 = W^T \Phi(X) + w_0
\]

where \( w_i \) is the weight; \( W \) is the weight vector; \( \phi(X) \) is the feature; \( \Phi(X) \) is the feature vector; and \( w_0 \) is the bias. In order to evaluate the prediction error, Vapnik [14] introduced a general error function, called the \( \varepsilon \)-insensitive loss function, defined by

\[
L_\varepsilon(d, f(X, W)) = \begin{cases} 0 & \text{if } |d - f(X, W)| \leq \varepsilon, \\ |d - f(X, W)| - \varepsilon & \text{otherwise} \\ \end{cases}
\]

Therefore, the penalty (loss) can be expressed by

\[
d_i - W^T \Phi(X) - w_0 - \varepsilon \leq \xi_i, i = 1, ..., Q \quad (3)
\]

\[
W^T \Phi(X) + w_0 - d_i \leq \xi_i, i = 1, ..., Q \quad (4)
\]

\[
\xi_i \geq 0, i = 1, ..., Q \quad (5)
\]

\[
\xi_i \geq 0, i = 1, ..., Q \quad (6)
\]

where \( \xi_i \) and \( \xi_i \) are non-negative slack variables used to measure the errors above and below the predicted function, respectively, for each data point. The empirical risk minimization problem can then be defined as [14, 15]

\[
\frac{1}{2} ||W||^2 + C \left( \sum_{i=1}^{Q} \xi_i + \sum_{i=1}^{Q} \xi_i \right)
\]

subject to the constraints in (3)-(6), where \( C \) is a user specified parameter for the trade-off between complexity and losses. To solve the optimization in (7), the Lagrangian in primal variables are constructed and its partial derivatives with respect to the primal variables have to vanish at the saddle point for optimality. Therefore, the simplified dual form then can be obtained as

\[
L_D(\Lambda, \Lambda') = \sum_{i=1}^{Q} d_i (\lambda_i - \lambda_i') - \varepsilon \sum_{i=1}^{Q} (\lambda_i + \lambda_i')
\]

maximize

\[
- \frac{1}{2} \sum_{i=1}^{Q} \sum_{j=1}^{Q} (\lambda_i - \lambda_i') (\lambda_j - \lambda_j') K(X_i, X_j)
\]

subject to

\[
\sum_{i=1}^{Q} (\lambda_i - \lambda_i') = 0
\]

\[
0 \leq \lambda_i \leq C, i = 1, ..., Q
\]

\[
0 \leq \lambda_i \leq C, i = 1, ..., Q
\]

where \( K(X_i, X_j) = \Phi(X_i) \cdot \Phi(X_j) \) is called the kernel function. With the Lagrangian optimization done, the optimal weight vectors can be obtained as follows

\[
\hat{W} = \sum_{i=1}^{Q} (\lambda_i - \lambda_i') \Phi(X_i) = \sum_{i=1}^{Q} (\lambda_i - \lambda_i') \Phi(X_i)
\]

where \( n \) is the number of support vectors, and the index \( k \) only runs over support vectors. Finally, the optimal bias can be obtained by exploiting the Karush–Kuhn–Tucker (KKT) conditions [16, 17], as follows

\[
\hat{w}_0 = \frac{1}{n_w} \sum_{i=1}^{Q} d_i - \sum_{i=1}^{Q} \beta_i K(X_i, X_i) - \varepsilon \text{sign}(\beta_i)
\]

where \( n_w \) is the number of unbounded support vectors with Lagrangian multipliers satisfying \( 0 < \lambda_i < C \) and \( \beta_i = \lambda_i - \lambda_i' \). Therefore, the approximate regression model can be obtained as follows:

\[
f(X, \hat{\lambda}, \hat{\lambda}') = \sum_{i=1}^{Q} (\lambda_i - \lambda_i') K(X_i, X) + \hat{w}_0
\]

B. The VIKOR Method

The VlseKriterijumska Optimizacija I Kompromisno Resenje (VIKOR) method [18] is one applicable technique that can be implemented within multiple criteria decision making (MCDM). The VIKOR method includes the following steps [19, 20]:

Step 1: Determine the best \( f_j^* \) and worst \( f_j \) values of all criterion functions \( j = 1, 2, ..., n \), where \( f_j \) is the value of the \( j \)-th criterion function for alternative \( A_i \).

\[
\max_{i=1} \min_{f_j} f_j = \max_i f_j^* \quad , \quad \min_{i=1} \max_{f_j} f_j = \min_i f_j^* , \quad \text{if the } j \text{-th function represents benefit;}
\]

\[
\max_{i=1} \max_{f_j} f_j = \max_i f_j^* \quad , \quad \min_{i=1} \min_{f_j} f_j = \min_i f_j^* , \quad \text{if the } j \text{-th function represents cost.}
\]

Step 2: Calculate the values \( S_i \) and \( R_i \) \( i = 1, ..., m \) by

\[
S_i = \sum_{j=1}^{m} w_j (f_j^* - f_j) , \quad R_i = \sum_{j=1}^{m} w_j (f_j - f_j^*)
\]

Step 3: Calculate the values \( Q_i \) \( i = 1, ..., m \) using

\[
Q_i = \nu \frac{S_i - S^*}{S^* - S} + (1 - \nu) \frac{R_i - R^*}{R^* - R},
\]

where \( S^* = \min_i S_i \) , \( S = \max_{i=1} S_i \) , \( R^* = \min_i R_i \) , \( R = \max_i R_i \), and \( \nu \) is the weight used for the strategy of the maximum group utility.

Step 4: Rank the alternatives, sorting the values \( S_i, R_i, Q_i \) in decreasing order.

Step 5: As a compromise solution, propose alternative \( A^{(3)} \), which is ranked best by the measure \( Q_i \) (minimum) if the following two conditions are satisfied:

C1. Condition for acceptable advantage:

\[
Q(A^{(2)}) - Q(A^{(1)}) \geq D Q ,
\]

where \( A^{(2)} \) is the alternative with the second position on the ranking list by \( Q_i \) and
DQ = 1/(m - 1).

C2. Acceptable stability in decision making:
Alternative $A^{(i)}$ also has to be ranked the best by $S_i$ and/or $R_i$.
If one of these two conditions is not satisfied, then a set of compromise solutions consisting of the following is proposed:
(1) Alternatives $A^{(1)}$ and $A^{(2)}$, if only condition C2 is not satisfied; or
(2) Alternatives $A^{(1)}$, $A^{(2)}$, …, $A^{(N)}$, if condition C1 is not satisfied, where $A^{(k)}$ is determined by

$$Q(A^{(k)}) - Q(A^{(1)}) < DQ$$

for maximum $k$.

C. Artificial Bee Colony Algorithm

Inspired by the intelligent foraging behavior of honey bee swarms in the natural world, Karaboga [21] developed a bee swarm algorithm, called the artificial bee colony (ABC) algorithm, whose general implementation steps are as follows [21-23]:

Step 1: Randomly generate an initial population consisting of $N_f$ feasible solutions where each solution $x_i = (x_{i1}, x_{i2}, ..., x_{in})$ is an $n$-dimensional vector.

Step 2: Evaluate the fitness of the initial solutions generated in Step 1.

Step 3: Each employed bee produces a candidate food position $v_i = (v_{i1}, v_{i2}, ..., v_{in})$ from the old one in its memory by

$$v_{ij} = x_{ij} + q(x_{ij} - x_{ij}'), \quad \forall i = 1,2, ..., N_f; \quad \forall j = 1,2, ..., n,$$

where $q \in [1,2, ..., N_f]$ is a randomly chosen index and $r_{ij}'$ is a random number in the range $(-1, 1)$.

Step 4: Evaluate the fitness of the candidate solutions created in Step 3. An employed bee memorizes the candidate food position $v_i = (v_{i1}, v_{i2}, ..., v_{in})$ if the fitness corresponding to the candidate food position is superior to the fitness of its old food position. Otherwise, the employed bee keeps the old food position in its memory.

Step 5: An onlooker chooses a food source with a probability calculated by

$$p_{ib} = \frac{fit_i}{\sum_{i=1}^{N_f} fit_i}, \quad \forall i = 1,2, ..., N_f,$$

where $p_{ib}$ is the probability that the $i$th food source will be chosen by an onlooker as the target to forage and $fit_i$ is the fitness of the $i$th food source.

Step 6: Each onlooker produces a modification of the position of the selected food source based on (19).

Step 7: Evaluate the fitness of the modified solutions made in Step 6. An onlooker memorizes the new position if the fitness corresponding to the modified solution is higher than that of its previous position.

Step 8: Memorize the position of the best food source found so far by the employed bees and onlookers.

Step 9: The employed bee abandons the food source $x_i = (x_{i1}, x_{i2}, ..., x_{in})$ and becomes a scout if it cannot improve the fitness of the corresponding food position in $C_{\text{limit}}$ search cycles.

Step 10: Each scout becomes an employed bee again and discovers a new food source based on

$$x_{ij} = x_{ij}^{\min} + s_{n}^{j}(x_{ij}^{\max} - x_{ij}^{\min}), \quad \forall j = 1,2, ..., n,$$

where $x_{ij}^{\min}$ and $x_{ij}^{\max}$ are the upper and lower bounds of the $j$th decision variable, respectively; and $s_{n}^{j}$ is a random number in the range $(0, 1)$.

Step 11: Repeat Steps 3 through 10 for $MCN$ cycles and designate the position of the memorized best food source as the final optimal solution.

III. PROPOSED APPROACH

This study proposes an integrated approach for solving multi-response parameter design problems based on SVR, VIKOR, and the ABC algorithm that is described in detail as follows:

Step 1: State the problem according to the objectives of the quality improvement project.

Step 2: Determine the key quality characteristics and their specification limits.

Step 3: Determine the major design/process parameters and their operational limits.

Step 4: Identify the important noise factors.

Step 5: Determine the number of experimental levels and the values for all the experimental levels for each control/noise factor.

Step 6: Select appropriate orthogonal arrays to arrange the control factors and the noise factors, and design an experimental layout.

Step 7: Conduct the experiment and collect experimental data.

Step 8: Normalize the quality characteristics values obtained along with the values of the major design/process parameters into a range of $-1$ to $1$.

Step 9: Randomly divide the normalized experimental data into training data and test data.

Step 10: Train and determine an appropriate SVR model for each quality characteristic to model the relationship between control factors and the quality characteristic.

Step 11: Determine the best and worst values for each quality characteristic in the VIKOR method.

Step 12: Explore the experimental ranges of the major design/process parameters using the ABC algorithm where the relationships between the design/process parameters and the quality characteristics are described using the SVR models constructed in Step 10, and the ranks of the solutions are obtained by ranking the $Q$ values introduced in the VIKOR method in ascending order, with rank 1 indicating the best solution. The $Q_i$ value for the $i$th solution in the ABC algorithm is calculated using (17). In addition, the $S_i$ and $R_i$ values in (17) are obtained by

$$S_i = \sum_{j=1}^{n} w_j d_j,$$

$$R_i = \sum_{j=1}^{n} w_j d_j.$$
$R_i = \max (w_j d_j)$, \hspace{1cm} (23)

where $w_j$ denotes the weight regarding the $j^{th}$ key quality characteristic, $n$ is the total number of key quality characteristics, and $d_j$ is calculated by

$$d_j = \begin{cases} f_j - y_j & \text{if } LSL_j \leq y_j \leq IV_{j \min} \\ f_j - f_j & \text{if } y_j < LSL_j \\ 0 & \text{if } y_j > IV_{j \min} \end{cases}$$

(for an LTB quality characteristic), \hspace{1cm} (24)

$$d_j = \begin{cases} f_j - y_j & \text{if } IV_{j \max} \leq y_j \leq USL_j \\ f_j - f_j & \text{if } y_j > USL_j \\ 0 & \text{if } y_j < IV_{j \max} \end{cases}$$

(for an STB quality characteristic), \hspace{1cm} (25)

$$d_j = \begin{cases} f_j - y_j & \text{if } TG_j \leq y_j \leq USL_j \\ f_j - f_j & \text{if } y_j < TG_j \\ 1 & \text{if } y_j \leq LSL_j \text{ or } y_j > USL_j \end{cases}$$

(for an NTB quality characteristic), \hspace{1cm} (26)

where $y_j$ is the de-normalized value for the $j^{th}$ key quality characteristic in the $i^{th}$ solution in the ABC algorithm.

Step 13: Obtain the (near) optimal settings of the major design/process parameters and conduct a confirmation experiment.

Step 14: If the confirmation result is unsatisfactory, repeat the entire procedure.

IV. CASE STUDY

Fig. 1(A) is an MR16 light-emitting diode (LED) lamp that is used in most fixtures designed for a traditional MR16 halogen lamp. In order to maximize the overall lighting performance of an MR16 LED lamp, the TIR lens, shown in Fig. 1(B), requires an elaborate design. According to the objectives of the quality improvement project, five key quality characteristics that are crucial to downstream clients were determined through discussions with LED design engineers and quality managers as follows:

1. Luminous flux ($y_1$)
2. Viewing angle at 0 degrees ($y_2$)
3. Viewing angle at 45 degrees ($y_3$)
4. Viewing angle at 90 degrees ($y_4$)
5. Viewing angle at 135 degrees ($y_5$)

The specification limits, response types, and associated weights for the above five key quality characteristics are summarized in TABLE I. As a result of brainstorming with design engineers, one important material property and four main geometric parameters, as illustrated in Fig. 2, of a TIR lens were selected as control factors. They are as follows:

1. Lens material ($x_1$): the material used to fabricate the TIR lens.
2. Lens height ($x_2$): the height of the TIR lens.
3. Lens radius of curvature ($x_3$): the radius of curvature of the TIR lens.
4. Micro-lens diameter ($x_4$): the diameter of the micro-lens.
5. Micro-lens spacing ($x_5$): the spacing between two adjacent micro-lenses.

TABLE II summarizes the experimental settings for the design parameters and noise factors. Next, Taguchi L$_{18}$ ($2^1 \times 3^7$) and L$_{6p}$($3^4$) orthogonal arrays were selected as the inner and outer arrays to arrange the five design parameters and four noise factors, respectively. To conduct the experiment, the SolidWorks 2010 (http://www.solidworks.com) modeling software was used to construct a geometric model of the TIR lens according to the settings for design parameters $x_1$ to $x_5$ in TABLE III, as well as to build the geometric model of an LED emitter. The SolidWorks model constructed was then fed into TracePro 5.0 (http://www.lambdares.com) simulation software for it to carry out optical simulations with the parameter setting of the lens material ($x_1$). TABLE III presents a part of the collected experimental results.

Next, the LIBSVM 2.86 [24] software for the SVR technique was applied to construct the estimated mathematical models to describe the functional relationship between each quality characteristic and the five design parameters where the radial basis function (RBF) was used as the kernel function. Notably, the grid-search approach [24] was used to determine the best combination of parameters $C$, $\gamma$, and $\varepsilon$ for one problem. First, the values obtained for the five quality characteristics along with the values for the five design parameters in each experimental trial listed in TABLE III were normalized in the range -1 to 1. A five-fold cross-validation method was then applied to the normalized experimental data. TABLE IV summarizes the optimal parameters found in SVR and the mean squared errors (MSEs) for the selected approximation regression models for each quality characteristic.

To find the optimal settings for the five design parameters of the TIR lens, the ABC algorithm was used to explore the experimental ranges in which the functional relationships between design parameters and key quality characteristics were described via the SVR$_{1\gamma}$, SVR$_{2\gamma}$, SVR$_{3\gamma}$, SVR$_{4\gamma}$, and SVR$_{5\gamma}$ models. First, the best and worst values for each key quality characteristic from the VIKOR method were determined. For the first response, $f_{j1}^+$ and $f_{j1}^-$ were set as 216 and 202.5, respectively, based on TABLE I. In addition, the best and worst values for each of the remaining four NTB quality characteristics were set as $f_{j2}^\gamma = 80$, $f_{j3}^- = 75$, and $f_{j4,5}^\gamma = 85$ (for $j = 2, 3, 4, 5$) according to their specifications shown in TABLE I. The parameters $N_f$ and $C_{\lima}$ in the ABC algorithm were set as 10 and 50, respectively. The ABC
algorithm was coded in Visual C++ 6.0 and ran on a personal computer with an Intel Core 2 Quad 2.66 GHz CPU and 2 GB RAM. The algorithm was terminated when the best solution found so far could not be further improved over the last 50 search cycles. TABLE V summarizes the execution results of implementing the ABC search procedure for 10 runs where the asterisk denotes the selected optimal solution. The average and standard deviation for the CPU time were 101.3 (sec) and 38.3 (sec). In addition, the average and standard deviation for the fitness were 0.8988 and 0.0003. On the basis of the above information, the ABC algorithm can be considered an efficient and robust optimization method for finding the optimal settings for the design parameters of a product/process.

To verify the feasibility and effectiveness of the optimal parameter settings obtained for the TIR lens, a confirmation experiment using SolidWorks and TracePro software was conducted. The results are summarized in the first trial in TABLE VI. In addition, a Taguchi L9(3^4) orthogonal array was employed to design another nine confirmation trials, as shown by the second to tenth trials in TABLE VI, in order to evaluate the robustness of the optimal parameter settings. The simulation results in TABLE VI reveal that all five quality characteristics in all the trials confirm the specification requirements.

To demonstrate the superiority of the integrated approach proposed in this study over the Taguchi method for dealing with parameter design problems with multiple responses, the full experimental data, partially shown in TABLE III, were further analyzed using the Taguchi method. By examining the factors’ effects based on the S/N ratios along with the suggestions of engineers, the optimal settings for the design parameters were finally determined. Simulation experiments were then conducted and the results were summarized in TABLE VII. Based on TABLE VII, it can be seen that none of the ten experimental trials could provide a design for a TIR lens that makes all of the five quality characteristics fulfill the specification requirements. This implies that manually making some trade-off for considering all quality characteristics at one time is an undesirable and inappropriate method for resolving a multiple-response parameter design problem. This also provides adequate evidence that the ABC algorithm can search the SVR models in the entire experimental ranges, thus finding the (near) optimal settings for design parameters in continuous domains rather than in the limited discrete experiment levels. From the above results and analyses, the integrated approach proposed in this study can be considered for popularization as a feasible and effective tool for solving general multi-response parameter design problems in the real world.

### Fig. 1. An MR16 LED lamp and a TIR lens.
In this study, the SVR technique, VIKOR, and the ABC algorithm were applied to design an integrated procedure to deal with these complicated and troublesome problems. The feasibility and effectiveness of the proposed approach were demonstrated via a case study in which the design of a TIR lens used in fabricating an MR16 LED lamp was optimized. The experimental results indicate that the proposed solution procedure can provide highly robust design parameter settings for a TIR lens. A comparison revealed that the ABC algorithm can search the optimal settings of design parameters by exploring the SVR models in continuous domains, thus finding a more robust design. Therefore, the integrated approach proposed in this study can be considered feasible and effective and can therefore be popularized as a useful tool for resolving general multi-response parameter design problems in the real world.

V. CONCLUSIONS

The author would like to thank the National Science Council, Taiwan, ROC for supporting this research under Contract No. NSC 101-2221-E-159-009. He would also like to thank Raymond Huang for his invaluable assistance during this study.

REFERENCES


