# **Development of a particle swarm optimization based support** vector regression model for titanium dioxide band gap characterization

Taoreed O. Owolabi<sup>1, 2, 3, †</sup>

<sup>1</sup>Physics and Electronics Department, Adekunle Ajasin University, Akungba Akoko, Nigeria <sup>2</sup>Computer Information System Department, College of Computer Science and Information Technology, Imam Abdulrahman bin Faisal University, Dammam, Postal code 31433, Saudi Arabia

<sup>3</sup>Physics Department, King Fahd University of Petroleum and Minerals, Saudi Arabia

Abstract: Energy band gap of titanium dioxide (TiO<sub>2</sub>) semiconductor plays significant roles in many practical applications of the semiconductor and determines its appropriateness in technological and industrial applications such as UV absorption, pigment, photo-catalysis, pollution control systems and solar cells among others. Substitution of impurities into crystal lattice structure is the most commonly used method of tuning the band gap of TiO<sub>2</sub> for specific application and eventually leads to lattice distortion. This work utilizes the distortion in the lattice structure to estimate the band gap of doped TiO<sub>2</sub>, for the first time, through hybridization of a particle swarm optimization algorithm (PSO) with a support vector regression (SVR) algorithm for developing a PSO-SVR model. The precision and accuracy of the developed PSO-SVR model was further justified by applying the model for estimating the effect of cobalt-sulfur co-doping, nickel-iodine co-doping, tungsten and indium doping on the band gap of TiO<sub>2</sub> and excellent agreement with the experimentally reported values was achieved. Practical implementation of the proposed PSO-SVR model would further widen the applications of the semiconductor and reduce the experimental stress involved in band gap determination of TiO<sub>2</sub>.

Key words: band gap; lattice distortion; crystal lattice parameters; particle swarm optimization; support vector regression; titanium dioxide

Citation: T O Owolabi, Development of a particle swarm optimization based support vector regression model for titanium dioxide band gap characterization[J]. J. Semicond., 2019, 40(2), 022803. http://doi.org/10.1088/1674-4926/40/2/022803

#### 1. Introduction

Titanium dioxide (TiO<sub>2</sub>) is a wide band gap and n-type semiconductor that offers promising potential in several technological and industrial applications which include gas sensing, photo-catalysis, spintronic, optical interference coating and many others<sup>[1, 2]</sup>. This type of semiconductor has many interesting properties such as non-toxicity, a high level of photo-sensitivity, commercial accessibility and stability, which warrant experimental and theoretical manipulation of its band gap to meet the requirements of the desired application<sup>[3]</sup>. Other features that widen the applications of TiO<sub>2</sub> semiconductor include high refractive index, high dielectric permittivity, low extinction coefficient and super-hydrophilicity<sup>[4]</sup>. Wide applicability of the TiO<sub>2</sub> semiconductor is attributed to its tunable band gap since different applications require different band gaps which makes the need for band gap engineering paramount<sup>[2]</sup>. In photo-catalysis applications, contraction of the band gap is essential for ensuring absorption of solar light beyond the UV region which constitutes approximately 5% of the solar light. A wide band gap that characterizes pure TiO<sub>2</sub> semiconductor is disadvantageous in this application and band gap con-

©2019 Chinese Institute of Electronics

traction is therefore needed through a doping mechanism. Furthermore, the fast rate at which the photo-generated charges recombine, coupled with a slow rate of electron transfer makes band gap engineering of TiO<sub>2</sub> a necessity<sup>[5]</sup>. On the other hand, the wide band gap of the TiO<sub>2</sub> semiconductor attracts attention in spintronic application due to its promising ferromagnetism with high magnetic ordering temperature (Curie temperature)<sup>[6]</sup>. To this end, the need for a model that can effectively estimate and tune the band gap of TiO<sub>2</sub> semiconductor is of high importance so as to facilitate proper manipulation of the semiconductor band gap with a high degree of precision and eventually widen the semiconductor applicability. A large number of experimental research works have investigated the effects of several metals, nonmetals and even a combination of multiple dopants on the band gap of TiO<sub>2</sub> with the sole aim of manipulating its band gap for specific application. However, the search for doping materials that can effectively tune the band gap of TiO<sub>2</sub> to the desired value is still a subject of intense interest in the literature<sup>[4, 7–11]</sup>. The goal of this work is to develop a particle swarm optimization based support vector regression (PSO-SVR) model through which the band gap of TiO<sub>2</sub> semiconductor can be tuned to the desired value using crystal lattice distortion as the input to the model and further relieve experimental stress involved in band gap measurement. It is worth mentioning that this is the first time, to the best of author's knowledge, that the particle swarm optimization (PSO) algorithm would be hybridized with support vector

Correspondence to: T O Owolabi, taoreed.owolabi@aaua.edu.ng; owolabitaoreedolakunle@gmail.com Received 25 JULY 2018; Revised 1 OCTOBER 2018.

regression (SVR) for TiO<sub>2</sub> semiconductor band gap estimation.

The crystal structure of  $TiO_2$  is polymorphic in nature and the three most commonly studied phases are rutile, anatase and brookite. Creation of a space charge region through replacement of titanium lattice ions in  $TiO_2$  semiconductor with dopant ions has been reported as an effective way of controlling its band structure<sup>[12]</sup>. Doping induces distortion in the lattice structure of the semiconductor and alters the energy band gap. The relationship between the distortion and band gap for doped  $TiO_2$  semiconductor has not been studied in the literature using a computational intelligence based model and this work aims to cover this gap.

Support vector regression (SVR) is an intelligence learning tool that implements a kernel trick for non-linear problems<sup>[13]</sup>. It has proved effective in many practical applications since non-linearity characterizes many real life problems<sup>[14–18]</sup>. Thr predictive strength of SVR is governed by proper selection of its hyper-parameters which mainly include regularization factors, epsilon and kernel options of Gaussian or polynomial kernel functions<sup>[19]</sup>. Many methods have been developed for SVR hyper-parameter selection in the literature<sup>[20-23]</sup>. The particle swarm optimization (PSO) technique is implemented in this work for optimizing hyper-parameters of SVR as a result of the advantages of PSO over other optimization methods<sup>[24]</sup>. Among the merits of PSO is that its implementation is based on intelligence and does not involve calculation of overlapping and mutation. Furthermore, it has fast speed of research since it only allows transmission of information from most optimist particles. Hybridization of PSO and SVR leads to a robust model (PSO-SVR) through which band gap of doped TiO<sub>2</sub> semiconductors can be easily tuned and controlled.

The results of the simulation indicate that the developed PSO-SVR model can estimate the band gap of doped  $TiO_2$  semiconductors with root mean square error as low as 0.165 eV. Comparison of the estimated and experimental band gap also shows that the developed PSO-SVR model is promising and can effectively tune the band gap of the semiconductor, relieve the stress involved in the experimental approach and most importantly, widens the potential application of the semiconductors without loss of precision.

#### 2. Description of the proposed PSO-SVR model

The section presents the mathematical background of the developed hybrid model. The operational principles of SVR and PSO algorithms are presented.

#### 2.1. Support vector regression computational intelligence tool

Support vector regression (SVR) is a machine learning tool that was developed based on the structural risk minimization principle (extended from statistical learning theory) which represents a robust statistical machine learning methodology<sup>[25, 26]</sup>. Inclusion of  $\varepsilon$ -insensitive loss function leads to the development of SVR from support vector machine which is mainly used for classification purposes. In describing the mathematical formulation of SVR, we consider a training set of data  $(x_1, y_1) \dots (x_j, y_j) \in \mathbb{R}^N \times \mathbb{R}$  where  $x_j$  and  $y_j$  respectively represent descriptors and the band gap of doped TiO<sub>2</sub>. A linear regression model is constructed as defined in equation (1) in which w and b are the parameters to be determined. Parameters

er b assumes zero value for a dataset already preprocessed to have zero mean<sup>[18]</sup>

$$F(x) = \langle \omega, x \rangle + b \text{ with } \omega \in N \text{ and } b \in R, \tag{1}$$

where N represents the space of input patterns and  $\langle .,.\rangle$  denotes the dot product in  $N^{\rm [27]}$ 

SVR algorithm solves the optimization problem described by Eq. (2) as subjected to the constraints in Eq. (3)

$$\frac{1}{2} \|\omega\|^2 + C \sum_{j}^{N} \left(\xi_j + \xi_j^*\right), \tag{2}$$

where ||.|| represents the Euclidian norm

$$\begin{cases} y_j - \langle \omega, x \rangle - b \leqslant \varepsilon + \xi_j, \\ \langle \omega, x \rangle - y_j + b \leqslant \varepsilon + \xi_j^*, \\ \xi_j, \xi_j^* \ge 0. \end{cases}$$
(3)

The parameter *C* contained in Eq. (2) is referred to as the regularization or penalty factor. Its function is to penalize the training samples with deviations larger than the threshold (epsilon). The essence of the slack variables  $(\xi_j, \xi_j^*)$  is to maintain the objective of the algorithm in ensuring flat function in cases where some training points fall outside the  $\varepsilon$ -tube<sup>[28]</sup>. Meanwhile, training data points above, below and inside the  $\varepsilon$ -tube are respectively assigned slack variables  $\xi > 0$  and  $\xi^* = 0$ ,  $\xi = 0$  and  $\xi^* > 0$  and  $\xi^* = 0$ .

For non-linear regression problems like the one presented in this work, the training data-points were mapped to high dimension feature space using non-linear mapping function presented in Eq. (4) with kernel option  $\sigma$ .

$$\varphi(x_j, x_k) = \exp\left(\frac{-1}{2} \left(\frac{\|x_j - x_k\|}{\sigma}\right)^2\right).$$
(4)

The final regression problem is presented in Eq. (5) with Langrage multipliers  $\alpha$  and  $\alpha^*$ .

$$F(x,\alpha) = \sum_{j=1}^{N} \left(\alpha_j^* - \alpha_j\right) \varphi(x_j, x_k) + b.$$
(5)

In SVR algorithm, hyper-parameters  $C, \varepsilon$  and  $\sigma$  are to be properly chosen and selected in order to build a robust model. In this work, PSO was implemented for this optimization problem.

#### 2.2. Particle swarm optimization method of optimizing SVR hyper-parameters

PSO is a population based computational search algorithm which mimics the social behavior of fish schooling or birds flocking in its development, implementation and search for optimum solution. PSO has demonstrated its excellent optimization capability since it was proposed and has been extensively applied to problems in science, engineering as well as technological problems<sup>[24, 29, 30]</sup>. It involves swarms of particle in which each of the particles represents a potential solution. Particles of the swarm explore and exploit hyperspace in an iterative manner with cognitive and social reasoning abilities

T O Owolabi: Development of a particle swarm optimization based support vector .....

Tuble 1. Statistical analysis of the dataset.				
Parameter	Lattice parameter <i>a</i> (Å)	Lattice parameter <i>c</i> (Å)	Band gap (eV)	
Mean (eV)	3.783	9.498 928	2.943 27	
Maximum (eV)	3.807	9.667	3.458	
Minimum (eV)	3.763	9.363	1.75	
Standard deviation	0.008	0.042 441	0.356 839	
Correlation coefficient (%)	1.42	-42.80		

Table 1. Statistical analysis of the dataset

Table 2. Optimum model parameters.

SVR hyper parameter	Optimum value	
Regularization factor (C)	1	
Epsilon ( $\varepsilon$ )	0.054	
Kernel option ( $\sigma$ )	0.0203	
PSO optimum parameters	Optimum value	
Number of population	200	
Number of generation	50	

which include the identification of their own local best position and knowledge of the best position of their neighbors. Evolvement of the proposed solutions or particles proceeds after each iteration until an optimal solution is attained. Consider a given search space of possible solutions in which the position and velocity of a particle j in a time step t is represented by  $z_j(t)$  and  $v_j(t)$  respectively. The subsequent position  $z_j(t+1)$  of the particle is obtained by addition of its velocity  $v_j(t+1)$  with previous position as detailed in Eqs. (6) and (7)

$$z_{i}(t) = z_{i}(t-1) + z_{i}(t), \tag{6}$$

$$w_{j}(t) = w(t)v_{j}(t-1) + c_{1}r_{1}\left(z_{\text{Lbest}}(t-1) - z_{j}(t-1)\right) + c_{2}r_{2}\left(z_{\text{Gbest}}(t-1) - z_{j}(t-1)\right),$$
(7)

where  $c_1$  and  $c_2$  represents acceleration coefficients while  $r_1$  and  $r_2$  are random vectors.

While developing the PSO algorithm, the potential solution called a particle p(t) in a time step t is represented as N-dimensional vector in which the number of parameters to be optimized is equal to N. Among the terms used in describing and controlling the attainment of optimum solution in the PSO algorithm include initial weight, individual best and stopping criteria. The initial weight w(t) represents a parameter that controls and balances the learning rate between local exploitation and global exploration and decreases as the optimum solution is being approached. The individual best is the best position a particle attained as evaluated using the fitness function and subsequently updated when the fitness function of the present position is better than that of the previous position. The condition that halts the algorithm is called the stopping criteria. Two stopping criteria implemented in this work are that the algorithm should stop if the global best does not change over a specified number of generations or the number of generations attains a pre-defined value.

#### 3. Methodologies for model development

This section details the description of the employed dataset for modeling and simulation. The computational hybridization of the proposed two algorithms is also presented. The significance of the number of population and generation on the predictive strength of the developed hybrid model is investigated and presented in this section.

# 3.1. Dataset employed in developing the PSO-SVR model for TiO<sub>2</sub> band gap estimation

The hybrid PSO-SVR model was developed using 63 experimental crystal lattice constants and their corresponding band gaps. The lattice constants as well as band gaps were obtained from published work when different impurities and dopants were incorporated into lattice structure of pure  $TiO_2$  semiconductor<sup>[31–37]</sup>.

The average of the content of the dataset is revealed from the presented statistical results. The range of the dataset as well as the deviation is also presented in the table. This gives enough insight into the content of the dataset and limitation as well the predictive strength of the proposed PSO-SVR model can be inferred from the presented statistical outcome. The correlation coefficient between pair of parameters reveals the degree and the extent of linear relationship that exists between them. The outcomes of the statistical analysis presented in Table 1 show that distortion along the *c*-axis is negatively correlated with band gap while distortion along the a-axis of pure TiO<sub>2</sub> semiconductor is positively correlated with the band gap having a very weak linear relationship. From the results of the correlation, it shows that a linear model would perform poorly in modeling the relationship between lattice parameters and band gap of TiO<sub>2</sub> semiconductor and the need for a non-linear model becomes imperative. SVR has numerous advantages that warrant its utilization in this present work. Its predictive strength is not often affected by a small number of datasets and due to this feature it has been extensively applied in condensed matter physics where significant inference needs to be made from few experimental datasets due to the difficulty in obtaining experimental data<sup>[38-42]</sup>.

# 3.2. Description of the computational hybridization of PSO and SVR algorithm

Hybridization of PSO with SVR aims at optimizing SVR hyper-parameters using a PSO algorithm. The hybridization presented in this work was conducted within the MATLAB computing environment. The hyper-parameters that are to be optimized using PSO are the regularization factor (C), epsilon ( $\varepsilon$ ) and kernel option ( $\sigma$ ) of the Gaussian kernel function. The stages involved in the hybridization include data separation, particle initialization, evaluation of objective function using each of the particles, time updating, weight updating, velocity updating, position updating, updating of individual best as well as global best followed by the specified stopping criteria. This is a minimization problem in which the root mean square error (RMSE) between the estimated and experimental band gap serves as a guide for reaching the best solution. Each of the particles of the swarm consist of  $C\varepsilon$  and  $\sigma$  hyper-parameters and the incorporation of this parameters into SVR algorithm coupled with training dataset produces a model that would further be validated using a testing dataset. Evaluation of the RMSE using the testing dataset dictates the direction of the algorithm, whether to proceed to the next generation or initialize the stopping criteria. The optimum value of SVR hyperparameters as obtained from PSO optimization are presented



Fig. 1. (Color online) Sensitivity of the developed PSO-SVR model on the number of generation keeping the number of population at 200.

in Table 2. A stepwise overall description of the hybridization is illustrated as follows:

**Step I:** Data separation: the dataset is partitioned randomly and systematically in the ratio of 80 to 20. 80% of the dataset is the training dataset while the remaining 20% serves as the testing dataset.

**Step II:** Particle initialization: *j*th particle  $(p_{j,n}(t=0))$  of *n*th dimension ( the dimension in our present case is three) is initialized randomly with uniform distribution within a specified range  $[p_{\max}, p_{\min}]$ . In the present modeling  $P_{\max} = 5000$ , 1 and 0 for regularization factor, epsilon and kernel option respectively while  $P_{\min} = -P_{\max}$ . In the same vein, velocities  $(v_{j,n}(t=0))$  of the particles were initialized  $[v_{\max}, v_{\min}]$ . The maximum attainable velocity of *j*th particle over *M* chosen interval is given by Eq. (8).

$$v_{j,\max} = \frac{\left(x_{j,\max} - x_{j,\min}\right)}{M},\tag{8}$$

where  $v_{j,\max} = -v_{j,\min}$ .

**Step III:** Evaluation of the objective function: SVR hyperparameters encoded in each particle were used to train the SVR algorithm using the aside training dataset. The support vectors emanating from the model training were used to test the predictive strength of the model with the aid of a testing dataset on the basis of RMSE. The position of a particle with minimum RMSE is designated  $z_j(t=0) = z_j^*(t=0)$  while the corresponding minimum RMSE is represented by  $J_j(t=0) = J_j^*(t=0)$ . If this position corresponds to global best then,  $z_j^*(t=0) = z_j^{**}(t=0)$  and  $J_j^*(t=0) = J_j^{**}(t=0) = J_{best}^{*}$ . Otherwise, updating is required.

**Step IV:** time updating: t = t + 1

**Step V:** weight updating: An adaptive weight described by  $w(t) = \alpha w(t-1)$  is implemented in this work where  $\alpha$  (usually < 1) represents a decrement constant that is close to 1.

**Step VI:** velocity updating: Velocity of the particle is updated in accordance to Eq. (7) with  $c_1 = c_2 = 2$  while  $r_1$  and  $r_2$  span in the range of [0, 1].

**Step VII:** position updating: Position of each of the particle is updated in accordance with Eq. (6)

Step VIII: updating of the individual best: Individual best



Fig. 2. Correlation cross-plot between the estimated and experimental band gap.

update follows Eq. (9)

$$\begin{cases} \text{if } J_j(t) < J_j^*(t), \text{ position is updated as:} \\ z_j^*(t) = z_j(t) \text{ and fitness function as: } J_j^*(t) = J_j(t) \qquad (9) \\ \text{otherwise, proceed without update} \end{cases}$$

**Step IX:** updating of the global best: the global best is updated in accordance to equation (10)

$$\begin{cases} \text{if } J_j(t) < J_j^{**}(t), \text{ position is updated as: } z_j^{**}(t) = \\ z_{\min}(t) \text{ and fitness function as: } J_j^{**}(t) = J_{\min}(t) \qquad (10) \\ \text{otherwise, proceed without update} \end{cases}$$

**Step X:** stopping criteria: The algorithm should stop if the specified number of generations is met or one of the earlier stipulated conditions is satisfied. Otherwise, the algorithm goes back to **step IV**:

### 3.3. Effect of number population on the performance of hybrid PSO-SVR model

The impact of varying the number of populations to the performance of the developed PSO-SVR model was investigated and the outcome is presented in Fig. 1. The number of generations at which the model shows no further improvement in performance was obtained as 50 and the number of populations that optimizes the model was searched for using this number of generations. The number of populations and generations that optimize the developed PSO-SVR model are presented in Table 2.

#### 4. Results and discussion

This section presents the results of the hybrid model. The outcomes of the developed hybrid model are also compared with experimentally measured band gaps. Effects of several dopants on the band gap of the semiconductor using the developed model are also discussed and presented.

#### 4.1. Evaluation of the predictive strength of the developed PSO-SVR model

A correlation cross-plot between the band gap estimated using the proposed PSO-SVR model and that of the experimentally reported values is shown in Fig. 2 for testing set of data. The data-points which represent the band gap of doped  $TiO_2$ semiconductor show little disparity with reasonable degree of

Table 3. Performance evaluation parameters and their corresponding values.

Performance evaluation parameters	Training dataset	Testing dataset
RMSE (eV)	0.253	0.165
MAE (eV)	0.163	0.125
CC (%)	72.56	84.13



Fig. 3. (Color online) Effect of cobalt-sulfur co-doping on the band gap of  $TiO_2$  using the developed PSO-SVR model.

alignment. The value of the coefficient of correlation is presented in Table 3. The developed PSO-SVR model also demonstrates the low value of RMSE as well as the mean absolute error (MAE) as shown in Table 3. The presented results show that the developed model can effectively estimate the band gap of doped TiO<sub>2</sub> semiconductor with an average error of 0.125 eV.

# 4.2. Influence of cobalt-sulfur co-doping on the band gap of TiO<sub>2</sub> semiconductor using the developed PSO-SVR model

The predictive strength of the developed PSO-SVR model was further assessed using doped TiO<sub>2</sub>. The effect of incorporating cobalt and sulfur into the crystal lattice structure of TiO<sub>2</sub> semiconductor was investigated and the estimated band gaps were compared with the reported values<sup>[11]</sup>. The developed PSO-SVR model was only fed with the lattice parameters while the model utilized the support vectors generated during the training stage for its band gaps agree excellently with the experimental results and also show a similar trend of reduction as the concentration of the dopants increase<sup>[11]</sup>. These results further strengthens the potential of the model in widen the energy absorption capacity of the semiconductor.

# 4.3. Influence of nickel-iodine co-doping on the band gap of TiO<sub>2</sub> semiconductor using the developed PSO-SVR model

The ability of nickel-iodine co-doping in tuning the band gap of  $TiO_2$  semiconductor was also examined using the developed PSO-SVR model. The results of the model are compared with the experimental results in Fig. 4. The PSO-SVR estimated band gaps agree perfectly with the experimental values<sup>[43]</sup>. Both the computational intelligence model prediction (our present results) and experimental values show that



Fig. 4. (Color online) Effect of nickel-iodine co-doping on the band gap of  $TiO_2$  using the developed PSO-SVR model.



Fig. 5. (Color online) Effect of tungsten doping on the band gap of TiO<sub>2</sub> using the developed PSO-SVR model.

the nickel-iodine co-doping result in fluctuation of the band gap of the semiconductor and the band gap attains minimum value when the concentration reaches 10%.

# 4.4. Effect of tungsten doping on the band gap of TiO<sub>2</sub> semiconductor using the developed PSO-SVR model

The effect of metal doping (tungsten) on the band gap of TiO<sub>2</sub> semiconductor was also investigated using the developed PSO-SVR model. The result of the simulation shows that tungsten metal slightly alters the band gap of the semiconductor. These results are compared with the experimentally reported band gaps<sup>[32]</sup>. Fig. 5 shows the comparison between the PSO-SVR estimated band gaps and the experimental results.

## 4.5. Effect of indium doping on the band gap of TiO<sub>2</sub> semiconductor using the developed PSO-SVR model

Finally, the developed model was implemented in investigating the effect of indium particles on the band gap of TiO<sub>2</sub> semiconductor by feeding the experimentally reported lattice parameters<sup>[44]</sup> of the indium doped TiO<sub>2</sub> semiconductor to the model. The estimated and experimental band gaps were compared in Fig. 6. The result of the modeling and simulation using PSO-SVR model shows that increase in the concentration of indium raises the band gap of the semiconductor and further lowers its absorption capacity.

T O Owolabi: Development of a particle swarm optimization based support vector .....



Fig. 6. (Color online) Effect of indium doping on the band gap of  $TiO_2$  using the developed PSO-SVR model.

#### 5. Conclusion

In this work, PSO is hybridized with SVR algorithm leading to the development of PSO-SVR model for band gaps estimation of doped TiO<sub>2</sub> semiconductor. The developed PSO-SVR model was validated using a testing set of data and minimum RMSE and MAE were obtained. The ability of the model to generalize and predict unseen datasets was further investigated. The developed PSO-SVR model was used to estimate the influence of several dopants on the band gap of TiO<sub>2</sub> semiconductor and excellent agreement with the experimentally reported band gaps was obtained. The influence of cobalt-sulfur co-doping, nickel-iodine co-doping, tungsten doping and indium doping on the band gap of TiO<sub>2</sub> semiconductor as estimated using the developed model shows excellent agreement with the experimental values. Based on the results presented in this work, the developed PSO-SVR model shows promising potential in photo-catalysis and other important applications where band gap engineering of TiO<sub>2</sub> semiconductor is essential and the model can be deployed for industrial and technological use so as to ease the stress in experimental determination of band gaps of doped TiO<sub>2</sub> semiconductor.

#### Acknowledgements

The support of King Fahd University of Petroleum and Minerals is acknowledged. The author also would like to thank Kabiru O. Akande (University of Edinburgh) for his valuable time spent to review the first draft of this manuscript.

#### Appendix

- List of abbreviations
- SVR: Support vector regression
- PSO: Particle swarm optimization
- TiO<sub>2</sub>: Titanium dioxide semiconductor
- $\left(\xi_{j},\xi_{j}^{*}\right)$ : Slack variables
- C: Regularization factor
- $\varepsilon$ : Epsilon hyper-parameter
- $\sigma$ : Kernel option of Gaussian kernel function
- $\varphi(x_i, x_k)$ : Kernel function
- $z_j(t)$ : Position of a particle *j* at iteration *t*
- $v_j(t)$ : Velocity of a particle *j* at iteration *t*

### References

- Kumar M, Gupta A K, Kumar D. Mg-doped TiO<sub>2</sub> thin films deposited by low cost technique for CO gas monitoring. Ceram Int, 2015, 42(1), 405
- [2] Diebold U. The surface science of titanium dioxide. Surf Sci Rep, 2003, 48(5–8), 53
- [3] Yacoubi B, Samet L, Bennaceur J, et al. Properties of transition metal doped-titania electrodes: Impact on efficiency of amorphous and nanocrystalline dye-sensitized solar cells. Mater Sci Semicond Process, 2015, 30, 361
- [4] Sobczyk-Guzenda A, Owczarek S, Szymanowski H, et al. Iron doped thin  $TiO_2$  films synthesized with the RF PECVD method. Ceram Int, 20115, 41(6): 7496 doi: 10.1016/j.mssp.2014.10.035.
- [5] Hamadanian M, Karimzadeh S, Jabbari V, et al. Synthesis of cysteine, cobalt and copper-doped TiO<sub>2</sub> nanophotocatalysts with excellent visible-light-induced photocatalytic activity. Mater Sci Semicond Process, 2016, 41, 168
- [6] Ahmed S A. Annealing effects on structure and magnetic properties of Mn-doped TiO<sub>2</sub>. J Magn Magn Mater, 2016, 402, 178
- [7] Kernazhitsky L, Shymanovska V, Gavrilko T, et al. Photoluminescence of Cr-doped  $TiO_2$  induced by intense UV laser excitation. J Lumin, 2015, 166, 253
- [8] Potlog T, Dumitriu P, Dobromir M, et al. Nb-doped TiO<sub>2</sub> thin films for photovoltaic applications. Mater Des, 2015, 85, 558
- [9] Arunachalam A, Dhanapandian S, Manoharan C, et al. Physical properties of Zn doped TiO<sub>2</sub> thin films with spray pyrolysis technique and its effects in antibacterial activity. Spectrochim. Acta -Part A Mol Biomol Spectrosc, 2015, 138, 105
- [10] Mollavali M, Falamaki C, Rohani S. Preparation of multiple-doped TiO<sub>2</sub> nanotube arrays with nitrogen, carbon and nickel with enhanced visible light photoelectrochemical activity via single-step anodization. Int J Hydrogen Energy, 2015, 40, 12239
- [11] Siddiqa A, Masih D, Anjum D, et al. Cobalt and sulfur co-doped nano-size TiO<sub>2</sub> for photodegradation of various dyes and phenol. J Environ Sci (China), 2015, 37, 100
- [12] McManamon C, O'Connell J, Delaney P, et al. A facile route to synthesis of S-doped TiO<sub>2</sub> nanoparticles for photocatalytic activity. J Mol Catal.A Chem, 2015, 406, 51
- [13] Kaneda Y, Mineno H. Sliding window-based support vector regression for predicting micrometeorological data. Expert Syst Appl, 216. 59: 217. doi: 10.1016/j.molcata.2015.05.002.
- [14] Akande K O, Owolabi T O, Olatunji S O, et al. A novel homogenous hybridization scheme for performance improvement of support vector machines regression in reservoir characterization. Appl Comput Intell Soft Comput, 2016, 2016, 1
- [15] Owolabi T O, Akande K O, Olatunji S O. Computational intelligence method of estimating solid- liquid interfacial energy of materials at their melting temperatures. J Intell Fuzzy Syst, 2016, 31, 519
- [16] Owolabi T O, Akande K O, Olatunji S O. Computational intelligence approach for estimating superconducting transition temperature of disordered MgB<sub>2</sub> superconductors using room temperature resistivity. Appl Comput Intell Soft Comput, 2016, 2016, 1709827
- [17] Owolabi T O, Akande K O, Olatunji S O. Estimation of average surface energies of transition metal nitrides using computational intelligence technique. Soft Comput, 2017, 21, 6175
- [18] Owolabi T O, Faiz M, Olatunji S O, et al. Computational intelligence method of determining the energy band gap of doped ZnO semiconductor. Mater Des, 2016, 101, 277
- [19] Suleiman M A, Owolabi T O, Adeyemo H B, et al. Modeling of autoignition temperature of organic energetic compounds using hybrid intelligent method. Process Saf Environ Prot, 2018, 120, 79
- [20] Ghorbani M, Zargar G, Jazayeri-Rad H. Prediction of asphaltene

#### T O Owolabi: Development of a particle swarm optimization based support vector .....

precipitation using support vector regression tuned with genetic algorithms. Petroleum, 2016, 2, 301

- [21] Zhou L, Lai K K, L Yu. Credit scoring using support vector machines with direct search for parameters selection. Soft Comput, 2009, 13(2), 149
- [22] Nieto P J G, Fernández J R A, Suárez V M G, et al. A hybrid PSO optimized SVM-based method for predicting of the cyanotoxin content from experimental cyanobacteria concentrations in the Trasona reservoir: A case study in Northern Spain. Appl Math Comput, 2015, 260, 170
- [23] Garcia Nieto P J, Garcia-Gonzalo E, Sanchez Lasheras F, et al. Hybrid PSO-SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability. Reliab Eng Syst Saf, 2015, 138, 219
- [24] Kennedy J, Eberhart R. Particle swarm optimization. IEEE International Conference on Particle swarm optimization, 1995, 4, 1942
- [25] Vapnik V. The nature of statistical learning theory. Springer, 1995
- [26] Zhang X, Wang P, Liang D, et al. A soft self-repairing for FBG sensor network in SHM system based on PSO-SVR model reconstruction. Opt Commun, 2015, 343, 38
- [27] Basak D, Pal S, Patranabis D C. Support vector regression. Neural Inf Process Lett Rev, 2007, 11(10)
- [28] Timoteo R D A, Silva L N, Cunha D C, et al. An approach using support vector regression for mobile location in cellular networks. Comput Networks, 2016, 95, 51
- [29] Nazari A, Sanjayan J G. Modelling of compressive strength of geopolymer paste, mortar and concrete by optimized support vector machine. Ceram Int, 2015, 41(9B), 12164
- [30] Abido M A. Optimal power flow using particle swarm optimization. Int J Electr Power Energy Syst, 2002, 24(7), 563
- [31] Khan W, Ahmad S, Hassan M M, et al. Structural phase analysis , band gap tuning and fluorescence properties of Co doped TiO<sub>2</sub> nanoparticles. Opt Mater (Amst), 2016, 38, 278
- [32] Oseghe E O, Ndungu P G, Jonnalagadda S B. Photocatalytic degradation of 4-chloro-2-methylphenoxyacetic acid using Wdoped TiO<sub>2</sub>. J Photochem Photobiol A Chem, 2015, 312, 96
- [33] Tahir M, Amin N S. Photocatalytic CO<sub>2</sub> reduction with H2 as reductant over copper and indium co-doped TiO<sub>2</sub> nanocatalysts in a monolith photoreactor. Appl Catal A Gen, 2015, 493, 90

- [34] Rangel-Vazquez I, et al. Synthesis and characterization of Sn doped  $TiO_2$  photocatalysts: Effect of Sn concentration on the textural properties and on the photocatalytic degradation of 2,4-dichlorophenoxyacetic acid. J Alloys Compd, 2015, 643(S1), S144
- [35] Lin Y H, Tseng T K, Chu H. Photo-catalytic degradation of dimethyl disulfide on S and metal-ions co-doped TiO<sub>2</sub> under visible-light irradiation. Appl Catal A Gen, 2014, 469, 221
- [36] Yu L, Yang X, He J, He Y, et al. One-step hydrothermal method to prepare nitrogen and lanthanum co-doped TiO<sub>2</sub> nanocrystals with exposed {001} facets and study on their photocatalytic activities in visible light. J Alloys Compd, 2015, 637, 308
- [37] Lei X F, Xue X X, Yang H. Preparation and characterization of Agdoped TiO<sub>2</sub> nanomaterials and their photocatalytic reduction of Cr(VI) under visible light. Appl Surf Sci, 2014, 321, 396
- [38] Owolabi T O, Akande K O, Olatunji S O. Development and validation of surface energies estimator (SEE) using computational intelligence technique. Comput Mater Sci, 2015, 101, 143
- [39] Majid A, Khan A, Javed G, et al. Lattice constant prediction of cubic and monoclinic perovskites using neural networks and support vector regression. Comput Mater Sci, 2010, 50(2), 363
- [40] Cai C Z, Wang G L, Wen Y F, et al. Superconducting transition temperature T c estimation for superconductors of the doped MgB<sub>2</sub> system using topological index via support vector regression. J Supercond Nov Magn, 2010, 23(5), 745
- [41] Cai C Z, Xiao T T, Tang J L, et al. Analysis of process parameters in the laser deposition of YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7</sub> superconducting films by using SVR. Phys C Supercond, 2013, 493, 100
- [42] Owolabi T O, Akande K O, Olatunji S O. Estimation of superconducting transition temperature T C for superconductors of the doped MgB<sub>2</sub> system from the crystal lattice parameters using support vector regression. J Supercond Nov Magn, 2014
- [43] Giannakas A E, Antonopoulou M, Deligiannakis Y, et al. Preparation, characterization of N-I co-doped TiO<sub>2</sub> and catalytic performance toward simultaneous Cr(VI) reduction and benzoic acid oxidation. Appl Catal B Environ, 2013, 140/141, 636
- [44] Tahir M, Amin N S. Indium-doped TiO<sub>2</sub> nanoparticles for photocatalytic CO<sub>2</sub> reduction with H<sub>2</sub>O vapors to CH<sub>4</sub>. Appl Catal B Environ, 2015, 162, 98