






Article

Tailoring the Energy Harvesting Capacity of Zinc Selenide Semiconductor Nanomaterial through Optical Band Gap Modeling Using Genetically Optimized Intelligent Method

Olusayo Olubosede ¹, Mohd Amiruddin Abd Rahman ^{2,*}, Abdullah Alqahtani ³, Miloud Souiyah ⁴, Mouftahou B. Latif ⁵, Wasiu Adeyemi Oke ⁶, Nahier Aldhafferi ³ and Taoreed O. Owolabi ^{2,7}

- ¹ Physics Department, Federal University Oye Ekiti, Oye Ekiti 371104, Ekiti, Nigeria; olusayo.olubosede@fuoye.edu.ng
- ² Department of Physics, Faculty of Science, Universiti Putra Malaysia (UPM), Serdang 43400, Malaysia; owolabitaoreedolakunle@gmail.com
- ³ Computer Information Systems Department, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, Dammam 31441, Saudi Arabia; Dr.abdullah1980@gmail.com (A.A.); naldhafferi@iau.edu.sa (N.A.)
- ⁴ Department of Mechanical Engineering, College of Engineering, University of Hafr Al Batin, P.O. Box 1803, Hafr Al Batin 31991, Saudi Arabia; souiyah@uhb.edu.sa
- ⁵ Instituut voor Kern- en Stralingsfysica Celestijnenlaan 200d–Bus 2418 B, 3001 Heverlee, Belgium; latifbak@gmail.com
- ⁶ Department of Mechanical and Mechatronics Engineering, Afe Babalola University Ado-Ekiti (ABUAD), P.M.B 5454, Ado-Ekiti 360231, Ekiti, Nigeria; wasiuad@abuad.edu.ng
- ⁷ Physics and Electronics Department, Adekunle Ajasin University, Akungba Akoko 342111, Ondo, Nigeria
- * Correspondence: mohdamir@upm.edu.my



Citation: Olubosede, O.; Abd Rahman, M.A.; Alqahtani, A.; Souiyah, M.; Latif, M.B.; Oke, W.A.; Aldhafferi, N.; Owolabi, T.O. Tailoring the Energy Harvesting Capacity of Zinc Selenide Semiconductor Nanomaterial through Optical Band Gap Modeling Using Genetically Optimized Intelligent Method. *Crystals* **2022**, *12*, 36. <https://doi.org/10.3390/cryst12010036>

Academic Editor: Ching-Hua Su

Received: 25 November 2021

Accepted: 22 December 2021

Published: 27 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Zinc selenide (ZnSe) nanomaterial is a binary semiconducting material with unique features, such as high chemical stability, high photosensitivity, low cost, great excitation binding energy, non-toxicity, and a tunable direct wide band gap. These characteristics contribute significantly to its wide usage as sensors, optical filters, photo-catalysts, optical recording materials, and photovoltaics, among others. The light energy harvesting capacity of this material can be enhanced and tailored to meet the required application demand through band gap tuning with compositional modulation, which influences the nano-structural size, as well as the crystal distortion of the semiconductor. This present work provides novel ways whereby the wide energy band gap of zinc selenide can be effectively modulated and tuned for light energy harvesting capacity enhancement by hybridizing a support vector regression algorithm (SVR) with a genetic algorithm (GA) for parameter combinatory optimization. The effectiveness of the SVR-GA model is compared with the stepwise regression (SPR)-based model using several performance evaluation metrics. The developed SVR-GA model outperforms the SPR model using the root mean square error metric, with a performance improvement of 33.68%, while a similar performance superiority is demonstrated by the SVR-GA model over the SPR using other performance metrics. The intelligent zinc selenide energy band gap modulation proposed in this work will facilitate the fabrication of zinc selenide-based sensors with enhanced light energy harvesting capacity at a reduced cost, with the circumvention of experimental stress.

Keywords: zinc selenide; support vector regression; lattice parameter; genetic algorithm; nano-particle size; energy band gap

1. Introduction

The synthesis and characterization of zinc selenide semiconductor nano-materials has attracted global attention lately due to the novel properties demonstrated by zinc selenide semiconductors compared to other members of chalcogenide groups [1–3]. The features contributing to the uniqueness of this class of semiconductors include the lower

electrical resistivity, non-toxicity, high transmission, insignificant lattice mismatching, wide energy band gap, as well as tunable light harvesting capacity over a range of spectrums [4]. These characteristic features, coupled with the quantum confinement effect, have enhanced and strengthened the practical application of zinc selenide semiconductors in various optoelectronics and photovoltaic applications, such as the thin-film window layer of solar cells, thin-film transistors, ultrasonic transducers, small integrated circuits, anti-reflection coating, energy generation, laser screens, light-emitting devices, and catalysis [5–7]. Tailoring the energy gap of zinc selenide semiconductor nanomaterials through doping lessens the loss of photons within the spectrum range, which contributes to charge carriers in photocurrents. The nano-structural size of the semiconductor and the crystal structural contraction (as well as elongation) emanated while doping are related to the energy gap in this contribution using an intelligent-based model.

The factors influencing the physical properties of zinc selenide semiconductors include the deposition parameters (for thin-film samples), experimental preparation conditions, synthesis techniques, and incorporation of dopants [7,8]. These factors ultimately control the crystal lattice parameters and the size of the semiconducting nanomaterial. Nickel particle incorporation into the lattice structure of zinc selenide semiconductors has been reported to excite photons, which, ultimately, promotes the carriers to the defect level with a consequential distortion in the lattice parameter at varying nano-sizes of the semiconductor [2]. This work aims to estimate the energy gap of doped zinc selenide semiconductors through the hybridization of a computational machine learning support vector regression algorithm and a genetic algorithm, using the lattice constant of the semiconductor and the size of the material as model inputs.

Support vector regression (SVR) belongs to the class of intelligent regression algorithms with high efficacy and efficiency in handling nonlinear real-life problems [9]. The algorithm was first proposed and developed by Vapnik, and effectively implements the kernel trick for dataset transformation into a space with characteristic high dimensionality [10]. The utilization of the kernel trick, convex optimization, and Lagrange multipliers enhances the global convergence of the algorithm, and subsequently averts local solutions. These unique features have significantly contributed to the wide application of this algorithm in different areas of study [11–14]. In an attempt to further enhance the robustness and reliability of SVR-based models, the user-defined parameters of the algorithm are optimized using a genetic algorithm (GA) in this work, with an evolutionary operational principle [15,16].

The remaining part of the work is organized as follows: Section 2 explains the mathematical formulation of a genetic algorithm and a support vector regression algorithm. The computational details and a description of the dataset employed for modeling are presented and described in Section 3 of the manuscript. The presentation and discussion of results are contained in Section 4, while Section 5 concludes the manuscript.

2. Description of Mathematical Formulation of the Employed Algorithms

The formulation of the employed support vector regression algorithm is presented in this section. A description of the hybridized genetic optimization algorithm is also presented in this section of the manuscript.

2.1. Support Vector Regression Mathematical Formulation

Consider the training samples $\{S, g^*\}$ of zinc selenide semiconductors doped with foreign materials and subsequently distort the lattice parameter at the nanoscale. The distorted lattice parameter and the semiconductor's nano-size are the descriptors to the model, and are represented as S , with a matrix size of $m \times D$, where m and D stand for the training sample data size and the sample dimension, respectively. The experimentally measured energy gaps of the semiconductor are represented as g^* , with a matrix size of $m \times 1$. The energy gaps to be predicted using the proposed hybrid model are empirically

represented by the function $f(S)$, with the following mathematical expression shown in Equation (1):

$$f(S) = \omega^T \sigma(S) + \gamma \quad (1)$$

where ω and γ represent the model vector to be optimized and the bias, respectively. Determination of $f(S)$ using the support vector regression (SVR) approach involves the solving and minimization of the optimization problem presented in Equation (2), with the aid of the constraints expressed in Equation (3) [10,17].

$$\frac{\omega^T \omega}{2} + \mu \sum_{j=1}^m (\eta_j + \eta_j^*) \quad (2)$$

$$\begin{cases} \omega^T \sigma(S_j) + \gamma - g^* \leq \varepsilon + \eta_j \\ g^* - \omega^T \sigma(S_j) - \gamma \leq \varepsilon + \eta_j^* \\ \eta_j^*, \eta_j \geq 0, j = 1, 2, 3, \dots, m \end{cases} \quad (3)$$

where μ = penalty factor, η = slack variable, η^* = slack variable, and ε = training sample error threshold. The value of μ penalizes and regularizes the deviation beyond the defined loss zone, while the incorporation of slack variables (η and η^*) becomes imperative to address the constraints that make the actualization of the training sample error threshold ε seem impossible [9,18]. μ further determines the equilibrium relationship existing between the function norm and empirical risk. Lagrange multipliers (ψ , ψ^* , χ and χ^*) are invoked to address the convex optimization problem with the Lagrange expression presented in Equation (4) [19].

$$\begin{cases} \omega^T \sigma(S_j) + \gamma - g^* \leq \varepsilon + \eta_j \\ g^* - \omega^T \sigma(S_j) - \gamma \leq \varepsilon + \eta_j^* \\ \eta_j^*, \eta_j \geq 0, j = 1, 2, 3, \dots, m \end{cases} \quad (4)$$

The partial derivative of Equation (4), with respect to ω , γ , η and η^* , leads to the expressions contained in Equation (5) [20–22].

$$\begin{cases} \omega = \sum_{j=1}^m (\psi_j - \psi_j^*) \sigma(S_j) \\ \mu - \psi_j - \chi_j = 0 \\ \mu - \psi_j^* - \chi_j^* = 0 \\ \sum_{j=1}^m (\psi_j - \psi_j^*) = 0 \end{cases} \quad (5)$$

The optimization problem is modified as presented in Equation (6), with the modified constraints presented in Equation (7).

$$\min \frac{1}{2} \sum_{j,i=1}^n (\psi_j^* - \psi_j) \zeta_{ij} (\psi_i^* - \psi_i) - \sum_{j=1}^m (\psi_j^* - \psi_j) + \varepsilon \sum_{j=1}^m (\psi_j^* + \psi_j) \quad (6)$$

$$\begin{cases} \sum_{j,i=1}^m (\psi_j^* - \psi_j) = 0 \\ \psi_j^* \geq 0, j = 1, 2, 3, \dots, m \\ \psi_j^* \leq \gamma, j = 1, 2, 3, \dots, m \end{cases} \quad (7)$$

where ζ_{ij} is the kernel function $\zeta_{ij} = \sigma(S_j)^T \sigma(S_i)^T$, with the expression provided in Equation (8) [20], as follows:

$$\zeta_{ij} = \exp(-\tau \|S_j - S_i\|^2) \quad (8)$$

where τ is the kernel parameter.

The final regression function is presented in Equation (9), as follows:

$$f(S) = \sum_{j=1}^m (\psi_j - \psi_j^*) \exp(-\tau \|S_j - S_i\|^2) + \gamma \quad (9)$$

The performance of the support vector regression-based model presented in Equation (9) can be significantly influenced through tuning of the model parameters, which include the regularization factor (μ), maximum error threshold epsilon (ϵ), and kernel parameter (τ) of the Gaussian kernel function [23]. The tuning of these parameters (μ, ϵ, τ) can be executed using a manual search, grid search, or meta-heuristic algorithms [24,25]. The choice of genetic algorithm for the optimization of these parameters is due to the ease of implementation of the genetic algorithm, convergence to global solution, and avoidance of premature convergence. These hyper-parameters are encoded as chromosomes in the genetic algorithm operational description using training and testing sets of data. All the genetic algorithm operators were implemented during the training and testing phases of model development using training and testing sets of data, respectively.

2.2. Genetic Algorithm

The genetic algorithm is a class of optimization algorithms that mimic and simulate naturally occurring evolutionary processes, with the aim of attaining optimum values for a set of parameters, through fitness selection from the initially defined population [15,26]. Genetic-based information is contained and encoded in a bit of string with a fixed length of characteristic features called individuals. The probable solutions are encoded in a chromosome-based architectural structure called individuals, in which the independent variables are represented by genes [27,28]. The primary operational principles of the algorithm include reproduction (new population generation), crossover (choosing probable parents and recombination), and mutation [29]. The commencement of an algorithm involves an initial population of potential individuals, while the subsequent generation through selection, crossover, and mutation improve the performance of the algorithm until global convergence is attained, which brings the algorithm to a stop [30]. Each of these basic operations of the genetic algorithm are controlled through probability setting. Several individuals are contained in the search space at the commencement of the algorithm iteration circle, while these individuals are evaluated in accordance with the defined problem-dependent fitness function [31,32]. The reproduction stage of the algorithm is often initiated randomly, followed by a parent selection process, which could be easily conducted through tournament selection or any other methods, such as the panmixia technique.

3. Computational Methodology

The computational development of SVR and GA is presented in this section. This section further presents the description of the employed set of data as well as the statistical analysis of the dataset.

3.1. Data Acquisition, Description and Statistical Analysis

Data samples from forty-three compounds of doped zinc selenide nanostructured semiconducting materials were employed in developing SVR-GA and SPR models for energy gap quantification. The data samples were extracted from different sources in the literature [4,5,7,8,33–38]. The descriptors such as the lattice parameter and the size of the semiconducting compounds were also extracted from the same sources with their corresponding energy gaps. Foreign material inclusion stretches the lattice parameters of the parent semiconductor and leads to distortion. Addition of dopants such as boron has been reported to widen the visible region transmittance due to the crystallinity of the semiconducting compound, which can be influenced by the size of nanoparticles [39]. Confinement of electrons and holes due to silver dopant incorporation into the parent zinc selenide semiconductor has also been reported to widen the energy gap of the materials

with increases in the crystalline size of the nanoparticles [40]. These are clear indications of the influence of lattice parameters and nano-crystalline size on the energy gaps of doped zinc selenide. In order to further investigate the potential of lattice parameters and the size of zinc selenide nanomaterials in estimating the energy gap of doped zinc selenide semiconductors, statistical analysis of the dataset was conducted. Table 1 presents the correlation coefficients between each of the descriptors and energy gaps; maximum, mean, minimum and standard deviations of the collected samples are also presented.

Table 1. Description of statistical analysis results.

Statistical Parameter	a (Å)	D (nm)	Energy Gap (ev)
Correlation coefficient	−0.569	−0.518	1.000
Maximum	5.935	82.200	3.790
Mean	5.601	21.546	2.795
Minimum	5.220	2.070	1.700
Standard deviation	0.153	20.219	0.506

In-depth statistical information about the entire dataset can be inferred from the values of mean, range (maximum and minimum), standard deviations and correlation coefficients.

3.2. Computational Incorporation of Genetic Algorithm in Support Vector Regression for Hyper-Parameter Optimization

The energy gap of zinc selenide nanostructured semiconductor was simulated and modeled through incorporation of genetic algorithm into support vector regression algorithm for the optimum selection of hyper-parameters, which include the penalty factor, kernel function and parameter as well as the epsilon (threshold modeling error on training samples). The simulation and modeling were conducted within the computing environment of MATLAB. The dataset available for pattern extraction was randomized and separated into training and testing set. The beauty of randomization process is that it promotes evenly dispersed and distributed intricacies needed for pattern acquisition between training and testing data samples in such a way that the model is not tested and validated on the information it has not acquired during training phase of model development. The procedures and processes undertaken for computational modeling of energy gap of doped zinc selenide are itemized below.

Step 1: Genetic algorithm initialization and initial parameter definition: Probable solutions were generated at this stage of model development and spanned across the defined search space for each of the support vector regression parameters. Penalty factor's upper and lower search space bounds were set as [1000, 1] while the space for the epsilon and kernel parameters was set as [0.09, 0.01] and [0.9, 0.1], respectively. The sensitivity (performance) of the developed hybrid model to the size of probable solutions within the search space was investigated purposely to set a balance between the exploration and exploitation strengths of the hybrid model for efficient prediction of energy gap of doped zinc selenide nanostructured semiconductor.

Step 2: Selection of the best chromosome among the probable ones on the basis of fitness function computation: This stage of model development incorporates the generated probable solutions from genetic algorithm into support vector regression algorithm. The potential of each of the chromosomes after being incorporated into SVR algorithm was determined using root mean square error of the testing set of data. The procedures for fitness function computation and determination are detailed as follows for each of the genetic evolutionary-based chromosomes:

- i. Each chromosome, which carries hyper-parameter information in a defined order, combines with a selected function (kernel) that helps in data transformation to space with higher dimension.
- ii. Combination of chromosomes, kernel function and hyper-parameter lambda ($E-7$) was implemented for training SVR algorithm using training set of samples.

The estimates of each of the trained SVR models during training phase (TP) were compared with the measured values using root mean square error (RMSE). The values of RMSE-TP for each trained model were ranked in ascending order. The support vectors corresponding to each of the trained models were also saved.

- iii. The saved support vectors for each of the trained models in Step 2 were employed for determining the energy gap of doped zinc selenide nanostructured semiconductors in the testing set of samples. The estimated energy gaps during testing phase (TSP) were compared with the experimental values using RMSE. The values of RMSE-TSP were recorded and ranked accordingly.
- iv. With the known value of RMSE-TSP for each of the trained and validated models, we could determine the best fit chromosome as characterized with lowest RMSE-TSP.

Step 3: Reproduction operator: This stage replaces the weak chromosomes from the initially generated probable solutions within the search space through application of reproduction operator. The probability for selection was set at 0.8 so as to ensure that the best-fit chromosomes identified in Step 2 are preserved through the subsequent stages of model development.

Step 4: Crossover operator deployment: For the purpose of establishing subsequences and portions exchanged between the parents and the offspring, crossover probability was implemented with a probability of 0.65. This stage allows weaker chromosomes (as determined by their fitness) to gradually become extinct while best-fit chromosomes are preserved.

Step 5: Population mutation: Randomization of chromosome portions was carried out at this stage. Mutation probability was set at 0.009.

Step 6: Control of the population size and stopping conditions of the hybrid algorithm: Step 2 to Step 5 were repeated while the chromosomes in the population were replaced with more fit individuals until similar value of RMSE-TSP was obtained for fifty consecutive iterations. The number of populations was varied from ten to hundred to investigate the influence of population size on both the convergence and performance sensitivity.

Step 7: Final model development with optimum parameters: Using the saved optimum parameters of genetic algorithm, final SVR-based model was developed while the values of both SVR and genetic algorithms that can enhance model reproducibility were saved. Figure 1 presents the details of the computational description of the developed hybrid model.

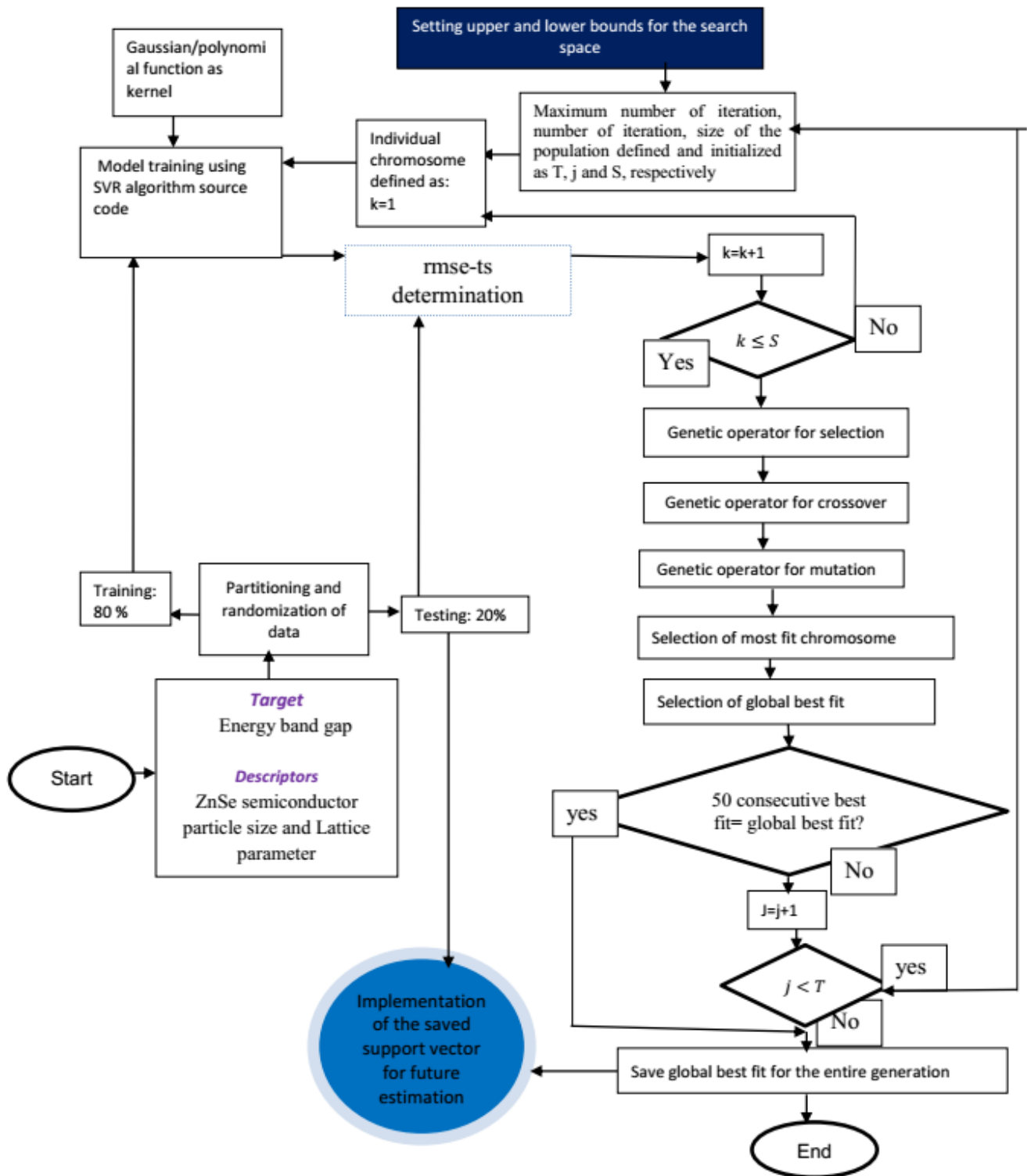


Figure 1. Description of the computational details of the developed hybrid model.

4. Results and Discussion

The results of the developed hybrid SVR-GA and SPR models are presented and discussed in this section. The influence of the hybrid model parameters and the comparison of the model performance are further discussed. The results of the developed hybrid models, when employed for external validation, are also presented.

4.1. Effect of Chromosomes Size on Penalty Factor, Epsilon and Kernel Parameter on Error Convergence of the Developed Hybrid Model

The convergence significance of the developed hybrid SVR-GA model on the model performance is presented in Figure 2. Figure 2a shows the variations in RMSE at different iterations for three different sizes of chromosomes in the model search space. The essence of investigating the significance of chromosome size on model performance is to strike a balance between model exploration and exploitation. As presented in Figure 2a, local convergence was observed, with ten chromosomes exploring the search space, while the points of local convergence are more prominent at the 7th and 20th iterations. When the number of chromosomes in the search space reached fifty, the global solution of the algorithm was attained. As can be observed in Figure 2a, the number of local convergences during the early iteration was reduced, while global convergence was achieved until the maximum iteration of one hundred was attained.

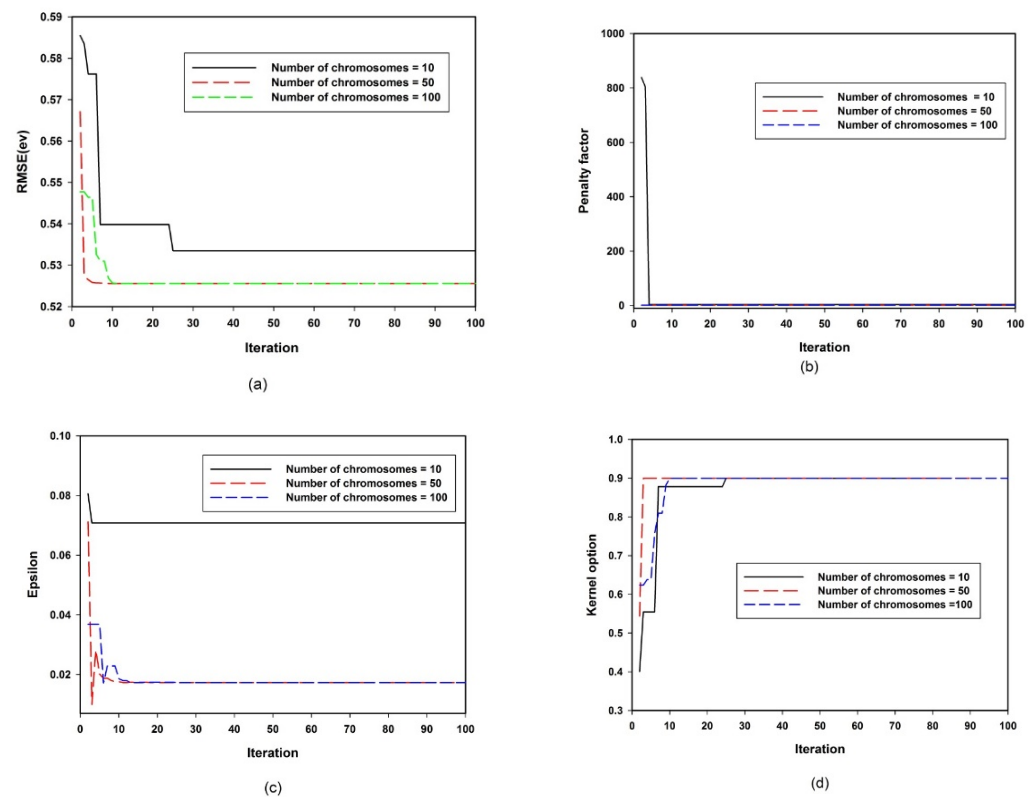


Figure 2. Dependence of chromosome population size on hybrid SVR-GA model convergence (a) effect of chromosome size on error convergence, (b) effect of chromosome size on the penalty factor, (c) effect of chromosome size on epsilon error threshold, (d) effect of chromosome size on Gaussian kernel parameter.

Similarly, the convergence of the penalty factor at different iterations for three categories of chromosome size is presented in Figure 2b. From the presented figure, it can be observed that the convergence of the measure of trade-off between the model complexity and the set maximum error threshold has an insignificant influence on model convergence. This demonstrates the robustness of the developed hybrid model. The influence of varying the size of the chromosomes present in the model search space on the defined maximum error threshold (epsilon) is presented in Figure 2c. This convergence shows a similar trend to the error (RMSE) convergence presented in Figure 2a. A lower number of chromosomes in the model search space results in a high value of epsilon, while an increase in the chromosome population size results in lower values of epsilon. Global convergence was achieved with fifty chromosomes in the search space. The kernel parameter for the

Gaussian kernel function is presented in Figure 2d at various chromosome population sizes. For each of the investigated chromosome numbers, global convergence was attained after twenty iterations. However, different numbers of chromosomes show different initial global convergences. Table 2 presents the obtained optimum value for each of the hybrid model parameters, as obtained after genetic algorithm optimization.

Table 2. Results of genetic algorithm for model optimization.

Model Parameter	Optimum Value
Penalty factor	1
Chromosomes number	50
Kernel parameter	0.9
Epsilon	0.0173
Kernel function	Gaussian
Hyper-parameter lambda	E-7

4.2. Performance Evaluation and Comparison for the Developed Hybrid SVR-GA and SPR Models

The obtained equation from the developed stepwise regression (SPR)-based model is presented in Equation (10), where a and D represent the lattice parameter in angstroms and the size of the zinc selenide nanoparticle, measured in nanometers, respectively. One of the merits of the SPR model over the SVR-GA model is the ease of implementation, while the developed SVR-GA model is superior when it comes to accuracy and precision.

$$SPR = 6.33104 - 0.59181a + 0.26606D - 0.04848aD \quad (10)$$

The performance of the developed hybrid SVR-GA-based model and the SPR model was evaluated and compared through computation of the correlation coefficient (CC), root mean square error (RMSE), and mean absolute error (MAE) during the training and testing stages of model development. Figure 3 presents the comparison of the two developed models using RMSE for the training and testing set of doped zinc selenide samples.

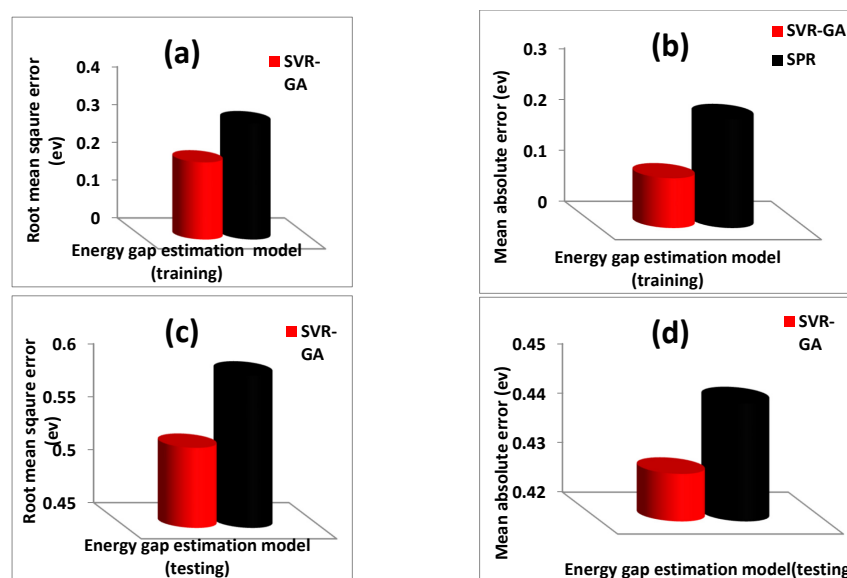


Figure 3. Performance of the developed energy gap estimation models. (a) Comparison of the performance of the developed SVR-GA and SPR models during training phase on the basis of RMSE, (b) comparison of the performance of the developed SVR-GA and SPR models during training phase on the basis of MAE, (c) comparison of the performance of the developed SVR-GA and SPR models during testing phase on the basis of RMSE, (d) comparison of the performance of the developed SVR-GA and SPR models during testing phase on the basis of MAE.

The developed hybrid SVR-GA model outperforms the SPR model, with a performance improvement of 33.68%, as presented in Figure 3a. The comparison of the two models, using the training MAE metric presented in Figure 3b, shows a 54.35% improvement of the SVR-GA model over the developed SPR model. The performance of the models was also compared while using the testing set of doped zinc selenide compounds. The outcomes of the comparisons are presented in Figure 3c,d, on the basis of RMSE and MAE, respectively. The developed SVR-GA model outperforms the SPR model, with a performance enhancement of 11.49% and 3.21%, on the basis of RMSE and MAE, respectively. The values for each of the computed performance measuring parameters, as well as the percentage improvement, are presented in Table 3.

Table 3. Computed performance of the developed model using evaluation parameters.

Dataset		SVR-GA (ev)	SPR (ev)	% Improvement
Training	CC	0.8988	0.7267	19.1440
	RMSE	0.2040	0.3076	33.6846
	MAE	0.0974	0.2134	54.3485
Testing	RMSE	0.5256	0.5939	11.4934
	MAE	0.4296	0.4439	3.2129

4.3. Energy Gap Tailoring in ZnSe Semiconductors with Copper–Indium Co-Doping Using Developed SVR-GA Model

The introduction of copper and indium nanoparticles, as co-dopants, into the lattice structure of the ZnSe semiconductor alters the energy gap of the parent semiconductor, as presented in Figure 4. The results of the developed SVR-GA model in this work align with the experimentally measured values [8]. Details of the experimental preparation of the samples can be found here [8]. Shifting of the absorption edge to longer wavelengths, as characterized by reduced energy gaps, was observed after copper–indium co-doping. Hence, the tendency to absorb photons of low energy was enhanced. The free carriers' concentration in the conduction band can be further influenced by the presence of the incorporated doped ions, which, ultimately, enhance the solar reflectance propensity of the parent semiconductor [8]. Since the estimates of the developed model agree excellently with the measured values, the photon absorption mechanism of the parent semiconductor can be enhanced for solar application using the developed model.

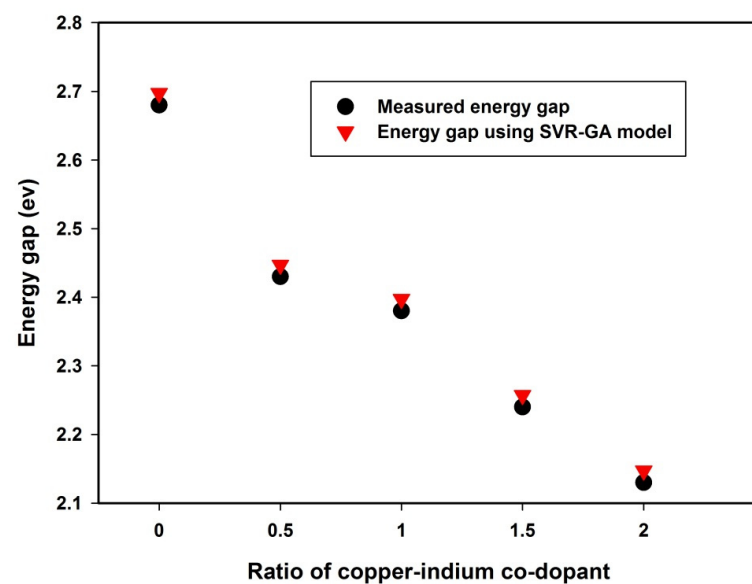


Figure 4. Effect of copper–indium co-doping (ratio) on energy gap of sulfur-based ZnSe.

4.4. Effect of Sulfur Particle Incorporation on Energy Gap of Doped ZnSe Semiconductor Using Developed SVR-GA Model

The energy gap tuning effect of sulfur particles' incorporation on the lattice structure of zinc selenide nanostructured semiconductors is presented in Figure 5, using the developed energy gap estimation model. The results of the developed model agree well with the measured values due to the intrinsic capacity of the SVR-based model to acquire and approximate the pattern relating to the semiconductor size, lattice parameters, and energy gap [36]. The energy gap of the parent semiconductor decreases with the increase in the concentration of sulfur nanoparticles, due to the absorption edge shifting to longer wavelengths. These observed characteristics of the doped zinc selenide semiconductor by the developed hybrid model (which have been validated by the experimental values) would be of great significance to tailor zinc selenide semiconductors for optoelectronics applications.

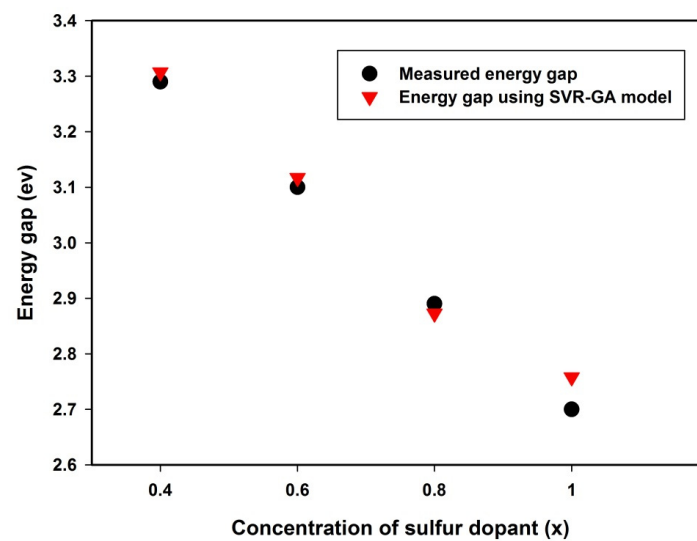


Figure 5. Influence of sulfur particles on energy gap of $\text{ZnS}_{1-x}\text{Se}_x$.

4.5. Outcomes of External Validation of the Developed Hybrid Model

The future generalization and potential of the developed hybrid SVR-GA model were further investigated and validated using another set of data extracted from different zinc selenide-based semiconductors, which were not included during the training and testing phases of model development.

Table 4 presents the results of the external validation. The estimates of the developed hybrid SVR-GA model were characterized and evaluated through CC, RMSE, and MAE computations, as presented in Table 4. The outstanding performance demonstrated by the developed hybrid model strongly indicates the potential of the developed model in tailoring zinc selenide-based nanostructured semiconductors in different applications, such as solar pigment, optoelectronics, among others.

Table 4. Results of further validation of the developed model.

a (Å)	D(nm)	Measured Energy Gap (ev)	SVR-GA (ev)	Residual
5.663	12.000	2.390 [2]	2.308	0.082
5.653	10.000	2.550 [2]	2.472	0.078
5.630	8.000	2.800 [2]	2.983	0.183
5.6373	18.791	2.676 [39]	2.765	0.089
5.6373	15.572	2.650 [39]	2.794	0.144
		2.630 [39]		
5.6373	14.083	Correlation coefficient = 97.76%, Root mean square error = 0.118 ev, Mean absolute error = 0.112 ev	2.725	0.095

5. Conclusions

The energy gap of the zinc selenide (ZnSe) nanostructured semiconductors with incorporated foreign materials is modeled through the stepwise regression (SPR) algorithm and the hybridization genetic algorithm (GA) in the support vector regression (SVR) algorithm. The lattice parameters and nano-size of the semiconductor were employed as the model's inputs. These inputs and their corresponding energy gaps were extracted from forty-three doped compounds of ZnSe nanostructured semiconductors for model development. The developed model was further validated using an external set of data extracted from another set of ZnSe compounds that were not included during model development. The two developed models (SPR and SVR-GA) were compared through CC, RMSE, and MAE evaluation, and the developed SVR-GA model shows outstanding performance over the SPR model, due to the excellent global convergence of GA, as well as the intrinsic features of the SVR-GA model to address non-linear problems through data transformation. The developed SVR-GA model investigates the influence of sulfur nano-particle incorporation on the energy gap of ZnSe semiconductors, as well as the energy gap tailoring effect of copper–indium co-dopants. The outstanding performance demonstrated by the developed hybrid model shows its potential in tailoring the energy gap of ZnSe semiconductors for solar, optoelectronic, and photocatalytic applications.

Author Contributions: Conceptualization, O.O. and M.A.A.R.; methodology, T.O.O.; software, A.A.; validation, M.S., M.B.L. and W.A.O.; formal analysis, N.A.; investigation, O.O.; resources, N.A.; data curation, M.A.A.R.; writing—original draft preparation, T.O.O.; writing—review and editing, M.B.L.; visualization, W.A.O.; supervision, M.A.A.R.; project administration, M.S.; funding acquisition, A.A. and T.O.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Universiti Putra Malaysia under Geran Putra IPM scheme with reference number of GP-IPM/2020/9694700. The APC was funded by the contributing authors.

Data Availability Statement: The raw data required to reproduce these findings are available in the cited references in Section 3.1 of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Suthar, D.; Chasta, G.; Patel, S.L.; Chander, S.; Kannan, M.D.; Dhaka, M.S. Impact of different annealing conditions on physical properties of ZnSe thin films for ecofriendly buffer layer applications. *Mater. Res. Bull.* **2020**, *132*, 110982. [[CrossRef](#)]
2. Divya, R.; Manikandan, N.; Vinitha, G. Synthesis and characterization of nickel doped zinc selenide nanospheres for nonlinear optical applications. *J. Alloys Compd.* **2019**, *791*, 601–612. [[CrossRef](#)]
3. Souri, D.; Reza, A.; Molaei, M.; Karimipour, M. ZnSe and copper-doped ZnSe nanocrystals (NCs): Optical absorbance and precise determination of energy band gap beside their exact optical transition type and Urbach energy. *Curr. Appl. Phys.* **2017**, *17*, 41–46. [[CrossRef](#)]
4. Balu, A.R.; Nagarethinam, V.S.; Basheer, M.G.S.; Thayumanavan, A.; Murali, K.R.; Sanjeeviraja, C.; Swaminathan, V.; Jayachandran, M. Influence of thickness on the microstructural, optoelectronic and morphological properties of nanocrystalline ZnSe thin films. *Mater. Sci. Eng. B* **2010**, *171*, 93–98. [[CrossRef](#)]
5. Ke, J.; Zhang, R.; Zhang, P.; Yu, R.; Cao, X. Investigation on structural and optical properties of ZnSe thin films prepared by selenization. *Superlattices Microstruct.* **2021**, *156*, 106965. [[CrossRef](#)]
6. Hien, N.T.; Yu, Y.Y.; Park, K.C.; Ca, N.X.; Chi, T.T.K.; Hien, B.T.T.; Thanh, L.D.; Do, V.; Tan, P.M.; Ha, P.T.T. Influence of Eu doping on the structural and optical properties of Zn 1-x Eu x Se quantum dots. *J. Phys. Chem. Solids* **2021**, *148*, 109729. [[CrossRef](#)]
7. Hasaneen, M.F.; Ali, H.M.; El-raheem, M.M.A.; Hakeem, A.M.A. Structure and optical properties of thermally evaporated Te doped ZnSe thin films. *Mater. Sci. Eng. B* **2020**, *262*, 114704. [[CrossRef](#)]
8. Zhang, T.; Wang, Y.; Pan, Z. Synthesis and characterization of Cu-/In-co-doped ZnS_xSe_{1-x} with tunable band-gaps as high near-infrared reflective color pigments. *Ceram. Int.* **2018**, *44*, 18851–18862. [[CrossRef](#)]
9. Science, N.; Phenomena, C.; Sabzekar, M.; Mohammad, S.; Hasheminejad, H. Chaos, Solitons and Fractals. *Chaos, Solitons and Fractals Interdiscip. J. Nonlinear Sci. Nonequilibrium Complex Phenom.* **2021**, *144*, 110738.
10. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 1995; pp. 30, 70–92.
11. Owolabi, T.O.; Amiruddin, M.; Rahman, A. Energy Band Gap Modeling of Doped Bismuth Ferrite Multifunctional Material Using Gravitational Search Algorithm Optimized Support Vector Regression. *Crystals* **2021**, *11*, 246. [[CrossRef](#)]

12. Akomolafe, O.; Owolabi, T.O.; Rahman, M.A.A.; Kechik, M.M.A.; Yasin, M.N.M.; Souiyah, M. Modeling superconducting critical temperature of 122-iron-based pnictide intermetallic superconductor using a hybrid intelligent computational method. *Materials* **2021**, *14*, 4604. [[CrossRef](#)] [[PubMed](#)]
13. Owolabi, T.O.; Abd Rahman, M.A. Modeling the optical properties of a polyvinyl alcohol-based composite using a particle swarm optimized support vector regression algorithm. *Polymers* **2021**, *13*, 2697. [[CrossRef](#)]
14. Shamsah, S.M.I.; Owolabi, T.O. Modeling the Maximum Magnetic Entropy Change of Doped Manganite Using a Grid Search-Based Extreme Learning Machine and Hybrid Gravitational Search-Based Support Vector Regression. *Crystals* **2020**, *10*, 310. [[CrossRef](#)]
15. Mitchell, M. Genetic Algorithms: An Overview. *Complexity* **1995**, *1*, 31–39. [[CrossRef](#)]
16. Zhou, L.; Lai, K.K.; Yu, L. Credit scoring using support vector machines with direct search for parameters selection. *Soft Comput.* **2009**, *13*, 149–155. [[CrossRef](#)]
17. Ju, X.; Liu, F.; Wang, L.; Lee, W.J. Wind farm layout optimization based on support vector regression guided genetic algorithm with consideration of participation among landowners. *Energy Convers. Manag.* **2019**, *196*, 1267–1281. [[CrossRef](#)]
18. Basak, D.; And, S.P.; Partababis, D.C. Support vector regression. *Neural Inf. Process.* **2007**, *11*, 203–224.
19. Lu, S.; Jiang, M.; Wang, X.; Yu, H.; Su, C. Optik Damage degree prediction method of CFRP structure based on fiber Bragg grating and epsilon-support vector regression. *Opt.-Int. J. Light Electron Opt.* **2019**, *180*, 244–253. [[CrossRef](#)]
20. Owolabi, T.O.; Saleh, T.A.; Olusayo, O.; Souiyah, M.; Oyenyin, O.E. Modeling the specific surface area of doped spinel ferrite nanomaterials using hybrid intelligent computational method. *J. Nanomater.* **2021**, *2021*. [[CrossRef](#)]
21. Owolabi, T.O. Determination of the Velocity of Detonation of Primary Explosives Using Genetically Optimized Support Vector Regression. *Propellants Explos. Pyrotech.* **2019**, *44*, 1282–1292. [[CrossRef](#)]
22. Owolabi, T.O. Modeling the magnetocaloric effect of manganite using hybrid genetic and support vector regression algorithms. *Phys. Lett. Sect. A Gen. At. Solid State Phys.* **2019**, *383*, 1782–1790. [[CrossRef](#)]
23. Owolabi, T.O. Development of a particle swarm optimization based support vector regression model for titanium dioxide band gap characterization. *J. Semicond.* **2019**, *40*. [[CrossRef](#)]
24. Owolabi, T.O.; Gondal, M.A. A hybrid intelligent scheme for estimating band gap of doped titanium dioxide semiconductor using crystal lattice distortion. *Comput. Mater. Sci.* **2017**, *137*, 249–256. [[CrossRef](#)]
25. Chamkalani, A.; Chamkalani, R.; Mohammadi, A.H. Hybrid of Two Heuristic Optimizations with LSSVM to Predict Refractive Index as Asphaltene Stability Identifier. *J. Dispers. Sci. Technol.* **2014**, *35*, 1041–1050. [[CrossRef](#)]
26. Tapia, A.; Nozal, A.R.; Reina, D.G.; Mill, P.; El, D.D.I.; El, D.D.I. Three-dimensional optimization of penstock layouts for micro-hydropower plants using genetic algorithms. *Appl. Energy* **2021**, *301*, 117499. [[CrossRef](#)]
27. Domashova, J.V.; Emtseva, S.S.; Fail, V.S.; Gridin, A.S.; Domashova, J.V.; Emtseva, S.S.; Fail, V.S.; Gridin, A.S. ScienceDirect Selecting an optimal architecture of neural network using genetic algorithm. *Procedia Comput. Sci.* **2021**, *190*, 263–273. [[CrossRef](#)]
28. Park, Y.; Yoo, J.; Park, H. A genetic algorithm for the vendor-managed inventory routing problem with lost sales. *Expert Syst. Appl.* **2016**, *53*, 149–159. [[CrossRef](#)]
29. Owolabi, T.O.; Amiruddin, M.; Rahman, A. Prediction of Band Gap Energy of Doped Graphitic Carbon Nitride Using Genetic Algorithm-Based Support Vector Regression and Extreme Learning Machine. *Symmetry* **2021**, *13*, 411. [[CrossRef](#)]
30. Ongcunaru, W.; Ongcunaru, P.; Janssens, G.K. Genetic algorithm for a delivery problem with mixed time windows. *Comput. Ind. Eng.* **2021**, *159*, 107478. [[CrossRef](#)]
31. Čičić, S.; Tomić, S. Genetic algorithm designed high efficiency laser power converters based on the vertical epitaxial heterostructure architecture. *Sol. Energy Mater. Sol. Cells* **2019**, *200*, 109878. [[CrossRef](#)]
32. Čičić, S.; Tomić, S. Automated design of multi junction solar cells by genetic approach: Reaching the 50% efficiency target. *Sol. Energy Mater. Sol. Cells* **2018**, *181*, 30–37. [[CrossRef](#)]
33. Kee, H.; Khoon, E.; Yao, K.; Mohd, N.; Adzir, M.; Pei, S.; Ahmad, M.; Abdul, A.; Huda, N.; Abidin, Z.; et al. Zinc selenide saturable absorber for ultrashort pulse fiber laser generation in C—band region. *Opt. Mater.* **2020**, *107*, 110100.
34. Shufeng, L.; Li, W.; Xueqiong, S.; Yong, P.; Dongwen, G.; Xiaowei, H. Optical properties of Co-doped ZnSe thin films synthesized by pulsed laser deposition. *Thin Solid Films* **2019**, *692*, 137599.
35. Khalfi, R.; Tounsi, A.; Merzouk, H. Effect of deposition time on structural and optical properties of ZnSe thin films grown by CBD method. *Opt. Mater.* **2020**, *106*, 109989. [[CrossRef](#)]
36. Sadekar, H.K.; Vithal, A.; Sharma, R. Bandgap engineering by substitution of S by Se in nanostructured ZnS_{1-x}Se_x thin films grown by soft chemical route for nontoxic optoelectronic device applications. *J. Alloys Compd.* **2011**, *509*, 5525–5531. [[CrossRef](#)]
37. Patil, N.M.; Nilange, S.G.; Yadav, A.A. Growth and characterization of ZnS_xSe_{1-x} thin films deposited by spray pyrolysis. *Thin Solid Films* **2018**, *664*, 19–26. [[CrossRef](#)]
38. Vs, G.K.; Mahesha, M.G. Characterization of spray deposited ternary ZnS_xSe_(1-x) thin films for solar cell buffers. *Surf. Interfaces* **2020**, *20*, 100509. [[CrossRef](#)]
39. Venkatachalam, S.; Kanno, Y.; Mangalaraj, D.; Narayandass, S.K. Effect of boron ion implantation on the structural, optical and electrical properties of ZnSe thin films. *Phys. B Condens. Matter* **2007**, *390*, 71–78. [[CrossRef](#)]
40. Yadav, K.; Jaggi, N. Effect of Ag doping on structural and optical properties of ZnSe nanophosphors. *Mater. Sci. Semicond. Process.* **2015**, *30*, 376–380. [[CrossRef](#)]