Estimation of Surface Energies of Transition Metal Carbides Using Machine Learning Approach

Taoreed O. Owolabi[1*], Kabiru O. Akande[2], and Sunday O. Olatunji[3]

[1] Physics Department, King Fahd University of Petroleum and Minerals, Dhahran, Kingdom of Saudi Arabia
[2] Electrical Engineering, King Fahd University of Petroleum and Minerals, Dhahran, Kingdom of Saudi Arabia
[3] Computer Science Department, University of Dammam, Dammam, Kingdom of Saudi Arabia

* Corresponding author. Tel. +966556202152; email: owolabitaoreedolakunle@gmail.com
Manuscript submitted October 02, 2014; accepted May 26, 2015.
doi: 10.17706/ijmse.2015.3.2.104-119

Abstract: Transition metal carbides (TMC) are characterized with high melting points which make experimental determination of their average surface energies a difficult task. A database of 3d, 4d and 5d TMC is hereby established using machine learning technique on the platform of support vector regression (SVR). SVR was built, trained and validated using some selected metals in periodic table and accuracy of 97.5% and 99.2% were achieved during training and testing phase respectively. Average surface energies of TMC were estimated using the trained and tested SVR model. Comparison of our results with surface energies from the first principle calculation and other theoretical results show agreement in terms of absolute values and the trend of variation depicted when Femi energy and density of state are analyzed using linear muffin-tin-orbital (LMTO). The computational ease of this approach in estimating average surface energies of TMC can be an edge over the existing theoretical methods.

Key words: Support vector regression, average surface energy, support vector regression model (SVRM) and transition metal carbide.

1. Introduction

The energy needed to create a new surface (surface energy) of materials is generally difficult to determine experimentally because of the need to heat the material to its melting phase [1]. Furthermore, TMC have high melting points which makes the experimental determination of their average surface energy more difficult than other materials with relatively low melting temperatures [2]. In conjunction with this difficulty, extrapolation of the surface tension to 0 kelvin also subjects the experimental results to certain level of inaccuracy [3]. However, TMC are technologically important as a result of their high melting point, hardness and excellent ability to conduct electricity which makes them fit well in layers of cutting tools, coating for solar application and conducting diffusion barriers in electronic devices to mention but few [2]. Hence, deep understanding of surface energy of these materials is paramount to unravel corrosion, oxidation, crystal growth and adsorption [4]. Calculation of surface energy from the first principle is known to be computationally demanding and is usually employed for few elements [5]. To the best of our knowledge, there is relatively few experimental data (with high degree of uncertainties) on surface energies of TMC and are rarely used for comparisons in literatures [2][6]. Consequently, the accuracy of our results is justified by comparison with generally accepted theoretical surface energies with respect to the absolute

values and trend or pattern of variation across TMC as depicted when analysis of Femi energy and density of states are carried out using LMTO [6]. The average surface energies obtained from the developed support vector regression model (SVRM) are comparable with the results obtained from several available models such as atomic sphere approximation using basis set of linear muffin-tin-orbital (LMTO-ASA) [6] ,bounding energy [6], bound-cutting [6],local density approximation (LDA) in density functional theory [2] and other standard theoretical methods . The simplicity of our model in terms of less computational complexity which saves time and resources makes huge difference.

SVR is one of the tools of artificial intelligence that is widely employed in several areas. In medical field, support vector machines help in identifying different kinds of cancers in their formative stages so as to take proper steps toward their cure [7-8]. Compressive strength of concrete that is difficult to determine is now being predicted using AI techniques [9]. AI techniques are not left out in predicting the properties of crude oil reservoirs [10][11]. Other areas where AI techniques are adopted include the estimation of atomic radii of elements [12], diagnosing mechanical fault [13], estimation of work function of semiconductors[14], estimation of surface energy of hexagonal close packed metals [15], assessing the thickness of metal plates [16], automatic recognition of off-line handwritten Arabic numbers [17], material characterization [18-19],prediction of superconducting transition temperature [20-21] among others [22]. The successes of support vector regression in solving several problems in material science coupled with the need to have accurate, reliable and easy means of estimating average surface energies of TMC motivated us to delve into this research work.

The empirical results of our simulations during the development of SVRM show that the developed model is capable of accurately estimating the average surface energy of TMC due to high accuracies of 97.5% and 99.2% achieved during training and testing phase respectively.

The remaining part of this work is organized as follows. Section 2 describes the proposed machine learning techniques (i.e. support vector regression) including the working principles of the SVR. Section 3 contains empirical studies that include the description of the dataset, computational methodology and the adopted strategies in searching for the optimum parameters. Section 4 presents and discusses results while section 5 states the conclusion and recommendation.

2. Proposed Method

This research work utilizes SVR derived from statistical learning theory [23-24] to develop a model (SVRM) through which average surface energies of TMC were estimated. SVR employs ε-insensitive loss function that controls the flatness of the generated pattern as well as the maximum tolerable deviations of the targets from the estimated values for all training dataset $(x_1, y_1),.....,(x_k, y_k)$ under the consideration with k number of samples [25]. Equation (1) represents a linear function in which $\langle w, x \rangle$ denotes the dot product in the space of $R^{'}$.

$$f(x, \alpha) = \langle w, x \rangle + b \tag{1}$$

where $w \in R^{'}$ and $b \in R$

For the purpose of establishing the goal of SVR in ensuring the flatness of equation (1), small value of w is desired through minimization of the Euclidean norm $\|w\|^2$ which makes the optimization problem of the regression looks like the one described in equation (2)

$$\text{minimize } \frac{1}{2}\|w\|^2$$
$$subject\ to \begin{Bmatrix} y_i - \langle w, x_i \rangle - b \le \varepsilon \\ \langle w, x_i \rangle + b - y_i \le \varepsilon \end{Bmatrix} \tag{2}$$

Existence of a function that is capable of providing error which is less than ε for all training pairs of the dataset is the condition under which equation (2) holds. The slack variables $(\xi_i \text{ and } \xi^*_i)$ are introduced in order to create room for another kind of error that may arise while dealing with real life problems. Therefore, equation (3) is modified and presented in equation (4).

$$\text{minimise } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{k}\left(\xi_i + \xi^*_i\right)$$
$$subject\ to \begin{Bmatrix} y_i - \langle w, x_i \rangle - b \le \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \le \varepsilon + \xi^*_i \\ \xi_i, \xi^*_i \ge 0 \quad \text{for all } i = 1, 2, \ldots\ldots, k \end{Bmatrix} \tag{3}$$

The optimization problem in equation (3) is solved in dual space representation. Meanwhile, Lagrangian multipliers $(\eta_i, \eta^*_i, \lambda_i \text{ and } \lambda^*_i)$ are invoked to transform the problem into dual space representation. Therefore, the Lagrangian for the equation (3) is presented in equation (4)

$$L = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{k}(\xi_i + \xi^*_i) - \sum_{i=1}^{k}\lambda_i\left(\varepsilon + \xi_i - y_i + \langle w, x_i \rangle + b\right)$$
$$-\sum_{i=1}^{k}\lambda^*_i\left(\varepsilon + \xi^*_i + y_i - \langle w, x_i \rangle - b\right) - \sum_{i=1}^{k}\left(\eta_i\xi_i + \eta^*_i\xi^*_i\right) \tag{4}$$

The saddle point of the Lagrangian function defined in equation (4) is easily located by equating the partial derivatives of the Lagrangian $(\text{with respect to } w, b, \xi_i \text{ and } \xi^*_i)$ to zero. These mathematical transformations give rise to the expressions presented in equations (5), (6) and (7).

$$w = \sum_{i=1}^{k}\left(\lambda^*_i - \lambda_i\right).x_i \tag{5}$$

$$\eta_i = C - \lambda_i \tag{6}$$

$$\eta^*_i = C - \lambda^*_i \tag{7}$$

The optimization equation is maximized by substituting equations (5-7) in (4) to give equation (8).

$$\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\lambda_i^* - \lambda_i\right)\left(\lambda_j^* - \lambda_j\right)\left(x_j, x_i\right) - \varepsilon\sum_{i=1}^{k}\left(\lambda_i^* + \lambda_i\right) + \sum_{i=1}^{k} y_i\left(\lambda_i^* - \lambda_i\right) = 0$$
$$subject\ \text{to}\ \sum_{i=1}^{k}\left(\lambda_i^* - \lambda_i\right) = 0,\ 0 \le \lambda^*{}_i\ \text{and}\ \lambda_i \le C \tag{8}$$

The Solutions (λ_i^* and λ_i) obtained from equation (8) were substituted in equation (1) and presented in equation (9)

$$f\ (x, \alpha) = \sum_{i=1}^{k}\left(\lambda_i^* - \lambda_i\right)\left\langle x_i, x\right\rangle + b \tag{9}$$

The concept of Kernel function is helpful in SVR algorithm for solving non-linear problems in which data is mapped into high dimensional feature space. The regression function in feature space can be written as shown in equation (10) with inclusion of the kernel function $K\left\langle x_i, x\right\rangle$

$$f\ (x, \alpha) = \sum_{i=1}^{k}\left(\lambda_i^* - \lambda_i\right)K\left\langle x_i, x\right\rangle + b \tag{10}$$

The Variables of kernel function control the structure of high dimensional feature space which measures the complexity of the final solution. Equations (11-14) describe most of the kernel functions that are obtainable in literatures [17]. They include Polynomial, Linear, Gaussian and Sigmoid functions as respectively described by Equations (11-14).

$$K(\overrightarrow{x_i}, \overrightarrow{x_j}) = (\overrightarrow{x_i} \cdot \overrightarrow{x_j} + 1)^d \tag{11}$$

$$K(x_i, x_j) = x_i^T x_j \tag{12}$$

$$K(\overrightarrow{x_i}, \overrightarrow{x_j}) = \exp\left(-\gamma\left\|\overrightarrow{x_i} - \overrightarrow{x_j}\right\|^d\right) \tag{13}$$

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r) \tag{14}$$

where $\gamma, r, and\ d$ represent kernel parameters

2.1. Working Principle of SVR

Principles of artificial intelligence are the basic principles adopt by SVR. It learns any pattern between descriptors and target so as to acquire a generalized pattern that guides its predictive ability during estimation of unknown target. The measures of inter-molecular energy of substances (cohesive energy) , the opposition offered by metals to uniform compression (bulk modulus) as well as lattice parameters of metals are the properties of material adopted as descriptors for developing SVR model through which average surface energies of TMC were estimated. The algorithm of SVR contains variables which are to be specified and varied by the user until desired performance is obtained from the model. These variables include the regularization factor which is also referred to as penalty factors that controls the trade-off between the flatness of the acquired pattern and the amount to which deviations larger than ε is allowed [25]. It is a penalty factor that has a wide limit of variation and controls the fitness of the model.

Epsilon, hyper-parameter and kernel option are among the variables of SVR that effect the performance of the model. Epsilon represents the maximum tolerable deviations of the estimated values from the target values. Hyper-parameter helps the model to select hyper-plane that minimizes the error while kernel option determines the structure of high dimensional feature space which controls the complexity of the developed model. In the case of the developed SVRM, the model was trained and tested before adopting it for estimation of average surface energies of TMC.

The training period of the model involves learning and acquisition of pattern (which may be complex) that needs to be generalized for future estimation of unknown target. The generalization of the acquired pattern during the training period of the model is achieved by the model in the course of comparing each generated target with the actual value so as to ensure generalized pattern. The accuracy, efficiency and the fitness of the generalized pattern can be validated through testing in which the model employs its acquired pattern during the training period to estimate unknown target with the aid of input descriptors. In the case of the developed SVRM, high accuracies of 97.5% and 99.2% were respectively obtained while training and testing the model. Well trained and tested SVRM was further used to estimate average surface energies of TMC.

2.2. Evaluation of the Generalization Performance of the Developed Model

The generalization performance of the developed model was evaluated using correlation coefficient (CC), root mean square error (RMSE) and absolute error (Ea). These parameters were respectively obtained through equation (15), (16) and (17).

$$cc = 1 - \left[\sum_{i=1}^{n} \frac{e_i}{E^2_{\exp}} \right] \tag{15}$$

$$rmse = \sqrt{\frac{1}{n} \sum_{i=1}^{n} e_i^2} \tag{16}$$

$$Ea = \sum_{i=1}^{n} |e_i| \tag{17}$$

where e_i , $E_{\exp}$ and n represent error (difference between the experimental and estimated average surface energy), experimental average surface energy and number of data point respectively.

3. Empirical Study

3.1. Description of the Dataset

The development of SVRM through which surface energy of TMC were estimated employs thirty-five experimental values of cohesive energy E_C, bulk modulus B and lattice parameters (a and c) of selected metals in periodic table. The descriptors and targets (average surface energies E_S) were drawn from the literatures [5], [27]–[33] and presented in Table 1.

Statistical analysis carried out on the dataset is depicted in Table 2a and 2b. The correlation coefficients presented in Table 2b show positive correlations for cohesive energy and bulk modulus while lattice parameters are negatively correlated with the average surface energies. Both positive and negative high correlations indicate a strong relationship between the chosen descriptors and the target which is best learned by support vector regression as indicated by the high percentage of accuracy obtained while developing the model.

Table 1. Dataset

Metals	E_S(eV)	E_C(eV)	a(A^0)	c(A^0)	B ($10^{11}Jm^{-3}$)
Li	0.525	1.630	3.491	3.491	0.116
Cs	0.095	0.804	6.045	6.045	0.020
Rb	0.110	0.852	5.585	5.585	0.031
K	0.130	0.934	5.225	5.225	0.032
Na	0.260	1.113	4.225	4.225	0.068
Nb	2.700	7.570	3.300	3.300	1.702
V	2.550	5.310	3.030	3.030	1.570
Cr	2.300	4.100	2.880	2.880	1.901
Mo	3.000	6.820	3.150	3.150	2.725
W	3.680	8.660	3.160	3.160	3.230
Fe	2.480	4.290	2.860	2.860	1.680
Ni	2.450	4.440	3.510	3.510	1.876
Cu	1.830	3.500	3.615	3.615	1.420
Be	2.700	3.320	2.270	3.590	1.144
Mg	0.760	1.520	3.210	5.210	0.369
Sc	1.275	3.900	3.310	5.270	0.435
Y	1.125	4.360	3.650	5.730	0.423
Ti	2.100	4.850	2.950	4.680	1.097
Zr	2.000	6.250	3.320	5.150	0.973
Hf	2.150	6.440	3.190	5.050	1.106
Tc	3.150	6.850	2.740	4.400	2.970
Re	3.600	8.030	2.760	4.460	3.715
Ru	3.050	6.740	2.710	4.280	3.152
Os	3.450	8.170	2.740	4.320	4.180
Co	2.550	4.390	2.510	4.070	1.948
Zn	0.990	1.350	2.660	4.950	0.804
Cd	0.740	1.160	2.980	5.620	0.621
Tl	0.575	1.880	3.460	5.520	0.382
Gd	1.110	4.140	3.630	5.780	0.405
Tb	1.130	4.050	3.600	5.700	0.399
Dy	1.140	3.040	3.590	5.650	0.411
Ho	1.150	3.140	3.580	5.620	0.397
Er	1.170	3.290	3.560	5.590	0.468
Tm	1.180	2.420	3.540	5.560	0.397
Lu	1.225	4.430	3.500	5.550	0.411

Table 2a. Statistical Analysis of the Dataset

	Mean	Median	Standard Deviation	Maximum	Minimum
E_S(eV)	1.73	1.28	1.07	3.68	0.095
E_C(eV)	4.11	4.1	2.3	8.66	0.804
a(A^0)	3.42	3.31	0.8	6.05	2.27
c (A^0)	4.62	4.95	1.01	6.05	2.86
B($10^{11}Jm^{-3}$)	1.22	0.8	1.15	4.18	0.02

Table 2b. Correlation between Each Pair of the Attributes of the Dataset

Attributes' pairs	Coefficient of correlation
E_C(eV)	0.903
a(A^0)	-0.684
c(A^0)	-0.652
B($10^{11}Jm^{-3}$)	0.925

Another dataset employed in estimating average surface energies of TMC using the developed model is presented in Table 3. It entails cohesive energies (Ec), lattice parameters (a) and bulk modulus (B) of 3d, 4d and 5d TMC obtained from literatures [30], [34]–[37]. The developed SVRM obtains average surface energies of TMC when the descriptors (presented in the Table 3) were fed into the model.

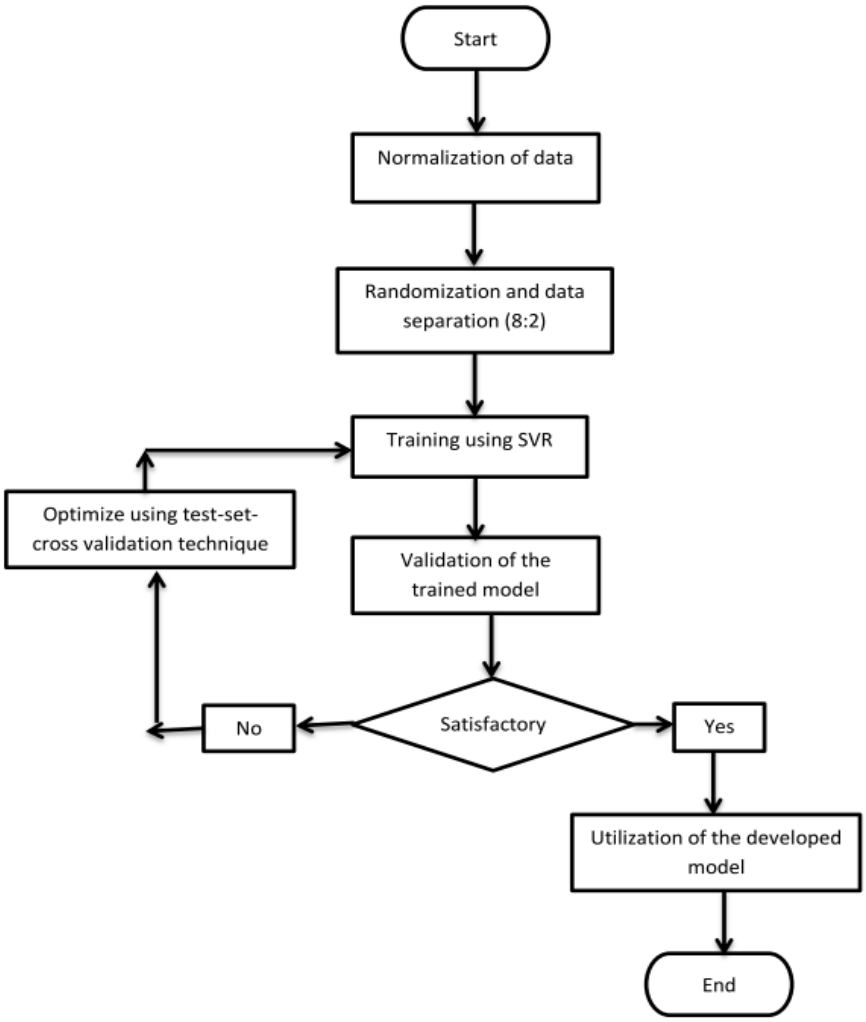
Table 3. Dataset for Estimating Average Surface Energy of TMC

TMC	Ec(eV)	a(A^0)	B(10^{11} Jm^{-3}))
ScC	6.37	4.72	4.68
TiC	7.16	4.33	5.26
VC	6.94	4.17	5.1
CrC	5.80	4.12	4.26
MnC	5.14	4.12	3.78
FeC	5.67	4.08	4.17
CoC	5.69	4.05	4.18
NiC	5.65	3.99	4.15
YC	6.39	5.19	1.243
ZrC	7.93	4.70	2.2
NbC	8.26	4.47	3.02
MoC	7.22	4.28	3.325
TcC	6.88	4.18	3.346
RuC	6.73	4.13	3.173
RhC	6.23	4.14	2.806
PdC	5.36	4.22	2.119
AgC	2.99	4.67	1.313
LaC	5.74	5.57	0.83
HfC	8.11	4.64	2.42
TaC	8.56	4.46	3.44
WC	8.24	4.26	3.96
ReC	7.47	4.17	4.04
OsC	7.36	4.12	3.92
IrC	6.84	4.13	3.28
PtC	6.34	4.20	2.58

3.2. Computational Methodology

This research work utilizes MATLAB computing environment for training and testing the SVR through which average surface energies of TMC were estimated. The MATLAB environment was also made used

while validating the developed model for determining average surface energies of TMC. The dataset for developing SVRM was normalized and reshuffled purposely to enhance efficient computations. The normalized dataset was further partitioned into training and testing phase in the ratio of 8 to 2 (which means that 80% of the thirty-five data-point was used to train the SVR while the remaining 20% was used to test the model). The developed SVRM (well trained and tested using SVR) was finally used to estimate average surface energies of TMC. Algorithm 1 shows the details of the computational methodology adopted while developing SVRM.



Algorithm 1: Computational methodology of the developed model

3.3. Strategy Adopted in Searching for Optimum Parameters

The accuracy, efficiency and the fitness of SVR depend greatly on the adopted strategies used in searching for optimum performance of the model. The developed SVRM performs optimally at certain values of regularization factor, hyper-parameter, kernel option and epsilon for a particular kernel function. We optimized each of these parameters using test-set-cross validation technique where the effect of one of the parameters (regularization factor, hyper-parameter, and epsilon and kernel option) on the performance of the model is determined while others are kept at constant values. The values of SVR parameters through which the developed model achieves its optimum performance are presented in Table 4.

Table 4. Optimum Parameters for the Developed SVRM

Parameters	Optimum value
C	1
Hyper-parameter(Lambda)	1E-7
Epsilon($\mathcal{E}$)	0.2
Kernel option	0.3
Kernel	Polynomial.

4. Results and Discussion

4.1. Development of SVRM

The development of SVRM that was employed in estimating average surface energies of TMC involves training and testing SVR using thirty-five values of experimental data of some periodic metals. The correlations between the experimental and estimated average surface energies in the course of training and testing the model are presented in Fig. 1 and Fig. 2 with correlation coefficients of 97.5% and 99.2% respectively shown in Table 5. The developed SVRM is characterized with high coefficient of correlation (cc), low root mean square error (rmse) and absolute error (Ea) as presented in Table 5. Since correlation shows the degree of similarity or closeness between two variables, these results mean that the average surface energies obtained from SVRM are 99% accurate and are very close to the experimental values. This further reflects the effectiveness and high accuracy of the developed model. Furthermore, the high values of correlation coefficients indicate excellent predictive and generalization ability of the developed SVRM which was further justified when the developed model was applied for the estimation of average surface energies of TMC.

Table 5. Determinants of the Quality of Performance of the Developed SVRM

	Training	Testing
cc	97.5%	99.20%
rmse	0.235	0.176
Ea	25.984	29.682

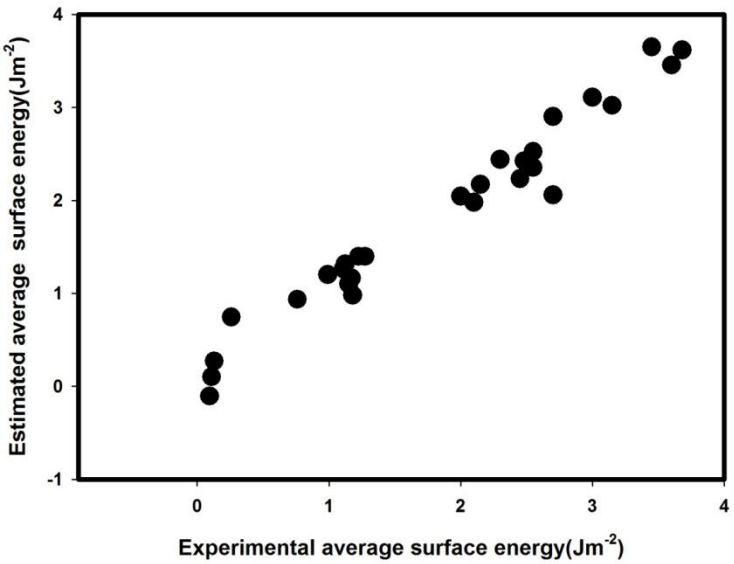


Fig. 1. Correlation between actual and estimated average surface energies while training SVRM

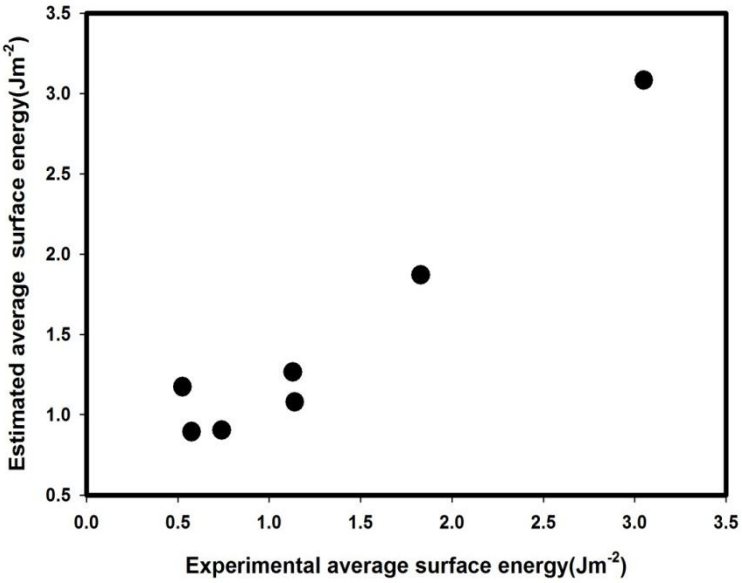


Fig. 2. Correlation between actual and estimated average surface energies while testing SVRM

4.2. Validation of the Developed SVRM through Estimation of Average Surface Energies of TMC

The developed SVRM was used to estimate average surface energies of 3d, 4d and 5d-TMC as detailed in the following sub-sections. Comparison of our results was made with the available surface energies obtained from the first principle calculations and several theoretical methods. Since the differences in the surface energies of two surfaces of materials are very important in surface phenomena, the similarity in trend of our results with the surface energies from other compared models is worthy of mentioning. Comparison of our results with several available surface energies from other theoretical models, in absolute values and trend is very crucial in justifying the accuracy of this approach.

4.2.1. Estimation of average surface energies of 3d-TMC

Table 5. Comparison between Average Surface Energies Obtained from the Developed SVRM (Our Results) with Atomic Sphere Approximation Using Basis Set of Linear Muffin-Tin-Orbital (LMTO-ASA)[6], Bounding Energy [6], Bound-Cutting [6], Local Density Approximation (LDA) in Density Functional Theory (DFT), Generalized Gradient Approximation (GGA) in DFT [2], Second Moment Approximation [2] and Broken Bond Model for 3d-TMC [2][34-35][38]. All in ev/atom.

3d-TMC	LMTO-ASA	SVR	LDA	Bounding energy	Bound cutting	GGA	Second-moment approximation	Broken bond model
ScC	0.67	1.04	0.56	0.86	0.68	0.47	0.79	0.57
TiC	0.83	1.03	0.69	0.98	0.84	0.54	0.89	0.65
VC	0.77	0.95	0.55	0.92	0.88	0.37	0.86	0.62
CrC	0.71	0.78	0.42	0.84	0.88	0.34	0.72	0.5
MnC	0.7	0.68	0.47	0.85	0.87	0.3	0.64	0.43
FeC	0.71	0.71	0.55	0.86	0.86	0.34	0.71	0.47
CoC	0.72	0.72	0.32	0.85	0.83	0.2	0.71	0.48
NiC	0.45	0.72	0.45	0.77	0.75	0.21	0.70	0.49

The developed SVRM was adopted to estimate average surface energies of 3d-TMC and the results are presented in Table 5. Table 5 also compares our results with other previously known surface energies and show agreement in terms of absolute values and trend obtained when Femi energy and density of states are analyzed [6]. Comparison of our results with surface energies from other theoretical models is presented in

Fig. 3. Average surface energy obtained from our simulations for stable 3d-TMC(that is CrC, MnC, FeC, CoC and NiC)[38] agree well with the results from other theoretical models. For example, surface energy obtained from our simulations (SVRM) for CrC, MnC, FeC ,CoC and NiC TMC are well predicted within theoretical results from atomic sphere approximation using basis set of linear muffin-tin-orbital (LMTO-ASA) [6] and second moment approximation[2] as indicated by the closeness in their absolute values and similar trend of variation.

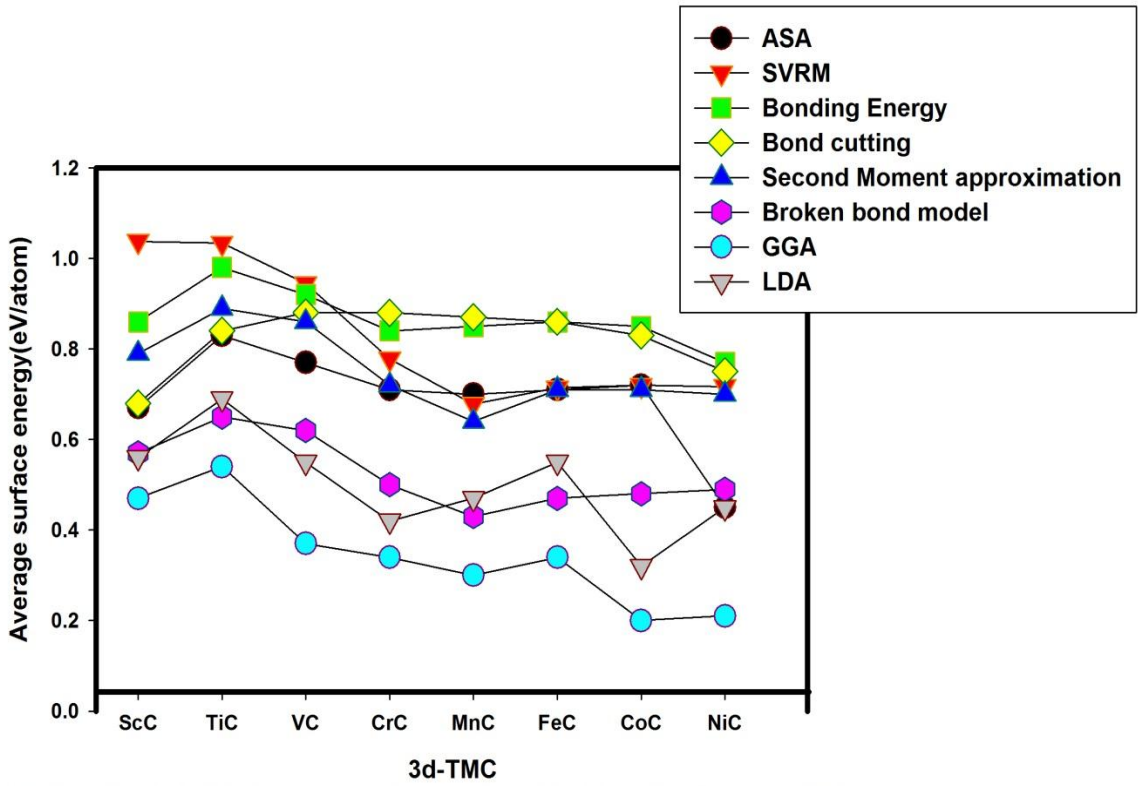


Fig. 3. Comparison between surface energies of 3d-transition metal carbides obtained from developed SVR model with other known models such as LMTO-ASA, bonding energy model, bond cutting model, second moment approximation, broken bound model, GGA and LDA.

4.2.2. Estimation of average surface energies of 4d-TMC

We further validated our developed SVRM through estimation of average surface energies of 4d-TMC and the results are presented in Table 6. For proper justification of the results obtained from our model, comparison of our results with surface energies from other theoretical methods was made and presented in Table 6 and Fig. 4. Absolute values of average surface energies obtained using the developed SVRM show excellent agreement with the results of bond cutting model derived from tight binding theory for both stable(YC, ZrC, Nbc, MoC and TcC)[34] and unstable(RuC, RhC, PdC and AgC) TMC [34]. TMC are characterized with different kinds of bounding properties such as metallic, ionic and covalent which makes bond cutting model that is mainly for metallically bonded material to be extended to TMC. Average surface energies of unstable TMC as obtained from our simulations show excellent agreement with surface energies obtained from second moment approximation [2] whose extension is mainly considered among the reference theoretical model in literature [2]. Results of our simulations are well predicted within other theoretical model as presented in Fig. 4.

Table 6. Comparison between Average Surface Energies Obtained from the Developed SVRM (Our Results) with Atomic Sphere Approximation Using Basis Set of Linear Muffin-tin-orbital (LMTO-ASA)[6], Bounding Energy[6], Bound-cutting[6], Local Density Approximation(LDA) in Density Functional Theory(DFT), Generalized Gradient Approximation(GGA) in DFT[2], Second Moment Approximation [2] and Broken Bond Model for 4d-TMC [2][38][34-35]. All in ev/atom.

4d-TMC	LMTO-ASA	SVR	LDA	Bounding energy	Bound cutting	GGA	Second-moment approximation	Broken bond model
YC	0.65	0.55	0.44	0.78	0.62	0.33	0.8	0.43
ZrC	0.86	0.80	0.66	0.94	0.82	0.53	0.99	0.64
NbC	0.87	0.89	0.69	0.87	0.9	0.49	1.03	0.72
MoC	0.77	0.86	0.64	0.77	0.9	0.38	0.9	0.66
TcC	0.69	0.86	0.61	0.77	0.88	0.37	0.86	0.66
RuC	0.69	0.86	0.45	0.77	0.86	0.31	0.84	0.68
RhC	0.68	0.82	0.41	0.75	0.82	0.38	0.77	0.62
PdC	0.47	0.68	0.48	0.65	0.73	0.32	0.67	0.52
AgC	0.39	0.44						

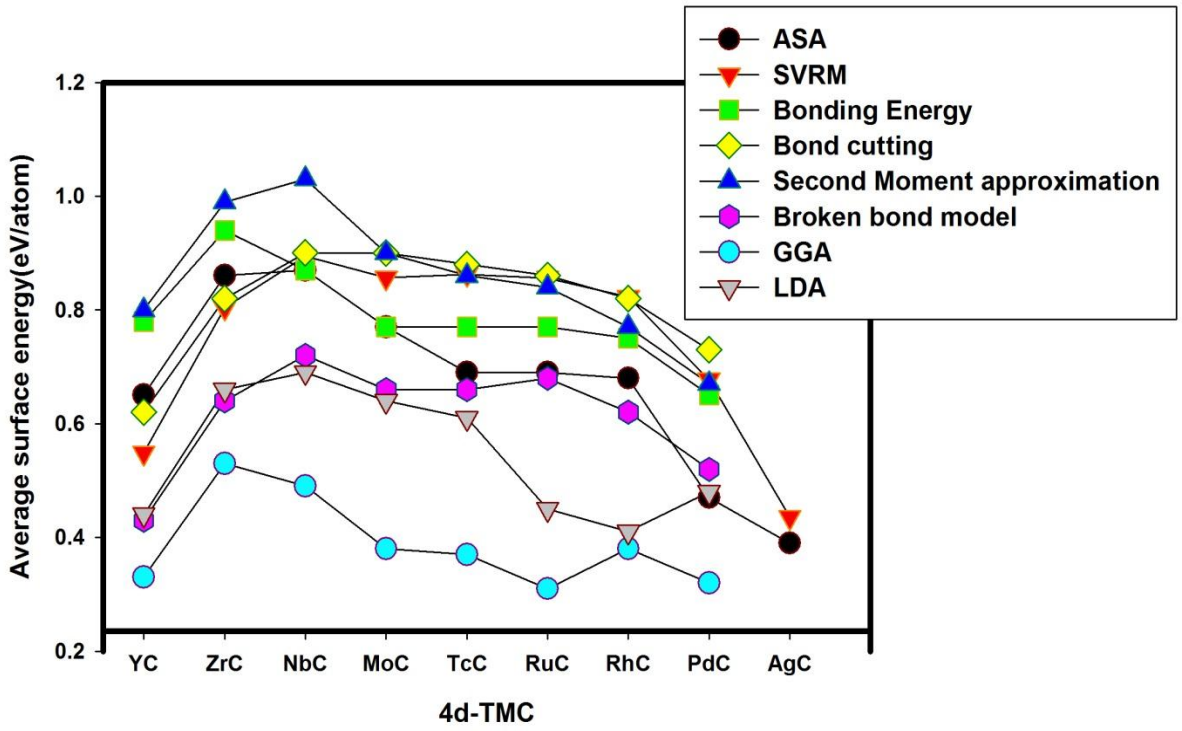


Fig. 4. Comparison between surface energies of 4d-transition metal carbides obtained from developed SVR model with other known models such as LMTO-ASA, bonding energy model, bond cutting model, second moment approximation, broken bound model, GGA and LDA.

4.2.3. Estimation of average surface energies of 5d-TMC

Average surface energies of 5d-TMC were obtained using the developed SVRM and the results are presented in Table 7. Average surface energies of 5d-TMC obtained from SVRM are well predicted within the values of surface energies obtained from other theoretical models under comparison as illustrated in Fig. 5. 5d-TMC that are proved stable (LaC, HfC, TaC and WC) in literature [35] are well predicted using SVRM while comparing the results of SVRM with surface energies from other models depicted in Fig. 5.

Table 7. Comparison between Surface Energies Obtained from the Developed SVRM (Our Results) with Atomic Sphere Approximation Using Basis Set of Linear Muffin-tin-orbital (LMTO-ASA)[6], Bounding Energy[6], Bound-cutting[6], Local Density Approximation(LDA) in Density Functional Theory(DFT), Generalized Gradient Approximation(GGA) in DFT[2], Second Moment Approximation[2] and Broken Bond Model for 5d-TMC [2][35][38-39]. All in ev/atom

5d-TMC	LMTOASA	SVM	LDA	Bounding energy	Bound cutting	GGA	Second-moment approximation	Broken bond model
LaC	0.7	0.47	0.33	0.81	0.63	0.25	0.71	0.39
HfC	0.9	0.82	0.7	0.97	0.82	0.52	1.01	0.68
TaC	0.88	0.91	0.68	0.89	0.87	0.52	1.07	0.7
WC	0.77	0.95	0.65	0.76	0.85	0.47	1.03	0.74
ReC	0.66	0.97	0.44	0.73	0.81	0.34	0.93	0.73
OsC	0.62	1.00	0.41	0.71	0.77	0.38	0.92	0.78
IrC	0.59	0.95	0.35	0.72	0.72	0.31	0.85	0.75
PtC	0.49	0.87	0.29	0.68	0.64	0.14	0.79	0.69

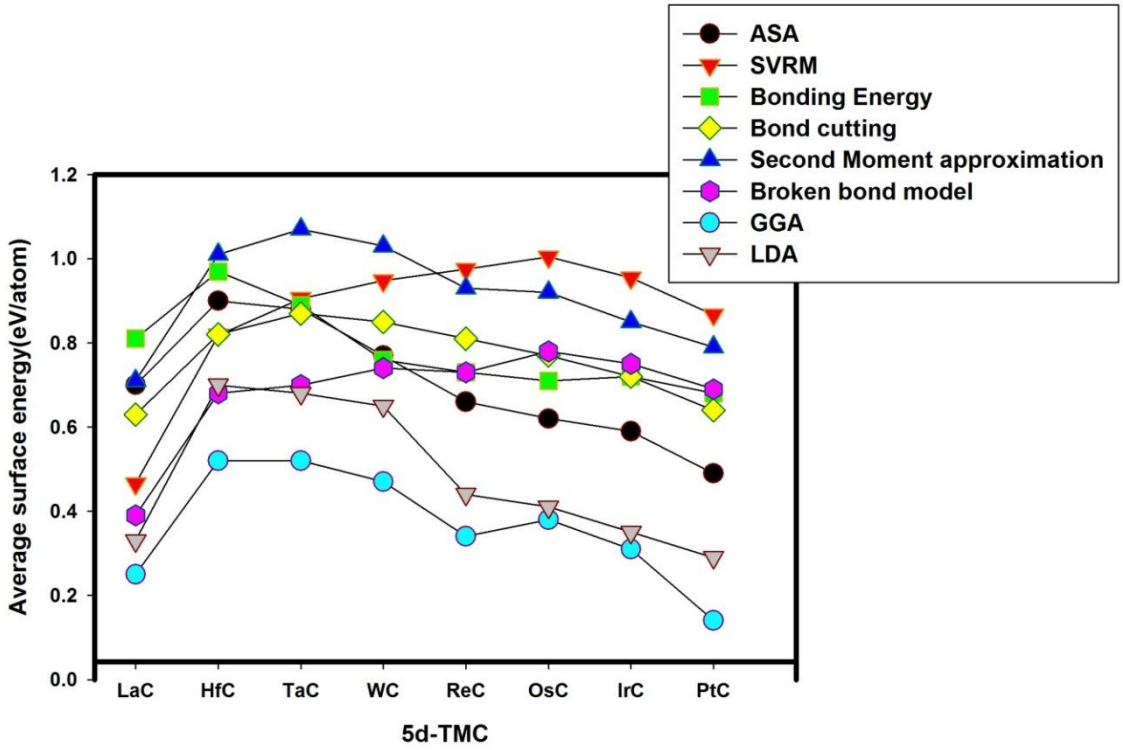


Fig. 5. Comparison between surface energies of 5d-transition metal carbides obtained from developed SVR model with other known models such as LMTO-ASA, bonding energy model, bond cutting model, second moment approximation, broken bound models, GGA and LAD.

5. Conclusion and Recommendation

We established a database for average surface energies of TMC using SVRM developed through training and testing SVR with properties of thirty-five selected metals from periodic table using test-set-cross validation optimization technique. Accuracy of the developed model is justified by comparison with surface energies from existing theoretical methods. Estimation of average surface energies from the developed

SVRM is therefore recommended due to its precision and easy computation which saves valuable time and resources.

References

[1] Tyson, W. R., (1997). Surface Free Energies of Solid metals: Estimation from liquid surface tension measurements. *Surf. Sci.*, *62*, 267–276.

[2] Liu, W., Liu, X., Zheng, W. T., & Jiang, Q. (2006), Surface energies of several ceramics with NaCl structure. *Surf. Sci.*, *600*, 257–264.

[3] Fu, B. Liu, W., & Li Z. (2010). Calculation of the surface energy of fcc-metals with the empirical electron surface model, *Appl. Surf. Sci.*, *256(22)*, 6899–6907.

[4] Aghemenloh, E., Yusuf, S., & Idiodi, J. O. A. (Mar. 2013). Surface energy calculation of hcp metals using the analytical equivalent crystal theory, *Mater. Chem. Phys.*, *138(2–3)*, 529–534.

[5] Fu, B. Q., Liu, W., & Li, Z. L. (Sep. 2009). Calculation of the surface energy of hcp-metals with the empirical electron theory. *Appl. Surf. Sci.*, *255(23)*, pp. 9348–9357.

[6] Hugosson, H. W., Eriksson, O., Jansson, U., & Ruban, A. V. (2004). Surface energies and work functions of the transition metal carbides. *557*, 243–254.

[7] Shini, M. A., Laufer S., & Rubinsky, B. (Sep. 2011). SVM for prostate cancer using electrical impedance measurements., *Inst. Phys.*, *32(9)*, 1373–1387.

[8] Cui, Y., Dy, J. G., Alexander, B., & Jiang, S. B. (2008). Fluoroscopic gating without implanted fiducial markers for lung cancer radiotherapy based on support vector machines., *Phys. Med. Biol.*, *53(16)*, N315–27.

[9] Akande, K. O., Owolabi, T. O., Twaha, S., & Olatunji, S. O. (2014). Performance comparison of SVM and ANN in predicting compressive strength of concrete. *IOSR J. Comput. Eng.*, *16(5)*, 88–94.

[10] Olatunji, S. O., Selamat, A., & Raheem, A. A. A. (Sep. 2011). Predicting correlations properties of crude oil systems using type-2 fuzzy logic systems, *Expert Syst. Appl.*, *38(9)*, 10911–10922.

[11] Olatunji, S. O., Selamat, A., & Abdul, R. A. A. (Jan. 2014). Improved sensitivity based linear learning method for permeability prediction of carbonate reservoir using interval type-2 fuzzy logic system, *Appl. Soft Comput.*, *14*, 144–155.

[12] Owolabi, T. O., Akande, K. O., & Olatunji, S. O. (2014). Estimation of the atomic radii of periodic elements using support vector machine, *Int. J. Adv. Inf. Sci. Technol.*, *28(28)*, 39–49.

[13] Baccarini, L. M. R., Rocha e Silva, V. V., de Menezes, B. R., & Caminhas, W. M. (June 2011). SVM practical industrial application for mechanical faults diagnostic, *Expert Syst. Appl.*, *38(6)*, 6980–6984.

[14] Owolabi, T. O., Akande, K. O., & Olatunji, S. O. (Sep. 2014). Support vector machines approach for estimating work function of semiconductors: Addressing the limitation of metallic plasma model, *Appl. Phys. Res., 6(5)*, 122–132.

[15] Owolabi, T. O., Akande, K. O., & Olatunji, S. O. (2015). Estimation of surface energies of hexagonal close packed metals using computational intelligence technique, *Appl. Soft Comput.*, *31*, 360–368.

[16] Ramos, H. G., Rocha, T., Král, J., Pasadas, D., & Ribeiro, A. L. (Aug. 2014). An SVM approach with electromagnetic methods to assess metal plate thickness, *Measurement*, *54*, 201–206.

[17] Mahmoud, S. A., & Olatunji, S. O. (2009). Automatic recognition of off-line handwritten Arabic (Indian) numerals using support vector and extreme learning machines. *Int. J. Imaging*, *2*.

[18] Swaddiwudhipong, S., Tho, K. K., Liu, Z. S., Hua, J., & Ooi, N. S. B. (Sep. 2005). Material characterization via least squares support vector machines. *Model. Simul. Mater. Sci. Eng.*, *13(6)*, 993–1004.

[19] Owolabi, T. O., Akande, K. O., & Olatunji, S. O. (Apr. 2015). Development and validation of surface energies estimator (SEE) using computational intelligence technique, *Comput. Mater. Sci.*, *101*, 143–

151.

[20] Owolabi, T. O., Akande, K. O., & Olatunji, S. O. (2014). Prediction of superconducting transition temperatures for Fe- based superconductors using support vector machine, *Adv. Phys. Theor. Appl.*, *35*, 12–26.

[21] Owolabi, T. O., Akande, K. O., & Olatunji, S. O. (Nov. 2014). Estimation of superconducting transition temperature T C for superconductors of the doped MgB2 system from the crystal lattice parameters using support vector regression. *J. Supercond. Nov. Magn.*.

[22] Akande, K. O., Owolabi, T. O., & Olatunji, S. O. (Jan. 2015). Investigating the effect of correlation-based feature selection on the performance of support vector machines in reservoir characterization. *J. Nat. Gas Sci. Eng.*, *22*, 515–522.

[23] Vapnik, V. (1995). *The Nature of Statistical Learning Theory Springer*.

[24] Kulkarni, S. R., & Harman, G. (2011). *Statistical Learning Theory : A Tutorial*, 1–25.

[25] Gupta, S. M. (2007). Support vector machines based modelling of concrete strength. *World Acad. Sci. Eng. Technol., 36*, 305–311.

[26] Olatunji, S. O. (2010). Comparison of extreme learning machines and support vector machines on premium and regular gasoline classification for arson and oil spill investigation, *ASIAN J. Eng. Sci. Technol.*, *1(1)*.

[27] Aghemenloh, E., Idiodi, J. O. A., & Azi, S. O. (Aug. 2009). Surface energies of hcp metals using equivalent crystal theory. *Comput. Mater. Sci.*, *46(2),* 524–530.

[28] Zhang, J. M., Wang, D. D., & Xu, K. W. (Dec. 2006). Calculation of the surface energy of hcp metals by using the modified embedded atom method. *Appl. Surf. Sci., 253(4)*, 2018–2024.

[29] Charles, K. (1976). *Introduction to Solid State Physics*.

[30] M. M., & Needs, R. J. (1991). Surface energy and stress of lead (111)and (110)surfaces, *Phys. Rev. B, 43(11)*, 8829–8833.

[31] Savino, E. J. (1992). Embedded-atom-method interatomic potentials for hcp metals, *Phys. Rev. B*, *45(22)*, 704–710.

[32] Igarashi, V. V. M., & Khantha, M. (1991). N-body interatomic potentials for hexagonal packed metal. *Philos. Mag. part B, 63(3)*, 1–10.

[33] Ni, A. K., Boer, F. R. D., Boom, R., Mattens, W. C. M., & Miedema, A. R. (1988). Cohesion in metals. *North-Holland, Amsterdam*.

[34] A. H., & Guillermet, G. G. A. F. (1992). Cohesive properties of 4d-transition-metal carbides and nitrides in the NaCl-type structure. *Phys. Rev. B*, 45(20), 557–567.

[35] Guillermet, A. F. (1993). Cohesive properties and electronic structure of 5d-transition-metal carbides and nitrides in the NaCl structure. *Phys. Rev. B*, *48(16)*.

[36] Korir, K. K., Amolo, G. O., Makau, N. W. Makau, & Joubert D. P. (2011). First-principle calculations of the bulk properties of 4 d transition metal carbides and nitrides in the rocksalt, zincblende and wurtzite structures, *Diam. Relat. Mater.*, *20(2)*, 157–164.

[37] L. F. Q., Li F. X., Cheng B., & Yin D. (1963). First principles study on the charge density and the bulk modulus of the transition metals and their carbides and nitrides, *Chinese Phys.*, *2287*.

[38] Guillermet, F. (1991). Band structure and cohesive properties of 3d-transition-metal carbides and nitrides with the NaCl-type structure, *Phys. Rev. B*, *43(18)*, 400–408.

[39] Guillermet, A. F., & Grimvall, G. (1992). Cohesive properties and vibrational entropyof 3d-transition metal carbide, *J.Phys. Chem. Solids*, *53(1)*.

Taoreed. O. Owolabi obtained B.Sc. degree in Physics and Electronics from Adekunle Ajasin University, Akungba-Akoko, Ondo State, Nigeria in 2010. He obtained his MS in physics from King Fahd University of Petroleum and Minerals, Saudi Arabia in 2014. He is currently doing his PhD (in Physics) in King Fahd University of Petroleum and Minerals, Saudi Arabia. His is also an Assistant lecturer in the Department of Physics and Electronics, Adekunle Ajasin University, Akungba –Akoko, Nigeria. His research interests include condensed matter physics and machine leaning.

Kabiru. O. Akande obtained B.Sc. degree in Electrical and Electronics Engineering from Ladoke Akintola University, Ogbomoso, Oyo State, Nigeria in 2010. He obtained his MS in Electrical Engineering from King Fahd University of Petroleum and Minerals, Saudi Arabia in 2014. He is currently doing his PhD in Electrical Engineering in King Fahd University of Petroleum and Minerals. Saudi Arabia. His research interests include digital Signal processing, blind equalization and machine leaning.

Sunday. 0. Olatunji is currently an Assistant Professor in the College of Computer Science & IT, University of Dammam, Saudi Arabia. He received the B.Sc. (Hons) Degree in Computer Science, Ondo State University (Now University of Ado Ekiti), Nigeria in 1999. He received M.Sc. Degree in Computer Science, University of Ibadan, Nigeria in 2003. He worked as a Lecturer in Computer Science Department, Ondo State University, Akungba Akoko, Nigeria, from 2001 to 2006, from where he proceeded to obtain another M.Sc. Degree in Information and Computer Science, King Fahd University of Petroleum and Minerals (KFUPM), Saudi Arabia (2006-2008). He obtained his PhD in Computer Science from the Universiti Teknologi Malaysia (UTM). He is a member of ACM and IEEE. He has participated in numerous research projects in and outside KFUPM including those with Saudi ARAMCO oil and gas Company, making use of artificial intelligence based methods in solving real industrial problems. His research interests include artificial intelligence and soft-computing, software engineering metrics prediction, machine learning and Data Mining