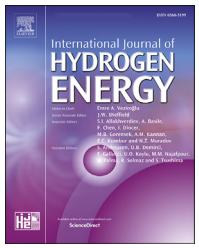
Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/he



A datamining approach to classify, select and predict the formation enthalpy for intermetallic compound hydrides

A. Djellouli [a], K. Benyelloul [b,c,*], H. Aourag [c], S. Bekhechi [d], A. Adjadj [a], Y. Bouhadda [b], O. ElKedim [e]

[a] Departement de Physique, Université de Tiaret, 14000, Algeria
[b] Unité de Recherche Appliquée en Energies Renouvelables, URAER, Centre de Développement des Energies Renouvelables, CDER, 47133, Ghardaïa, Algeria
[c] Departement de Physique, Université de Tlemcen, 13000, Algeria
[d] Departement de Physique Théorique, Université de Tlemcen, 13000, Algeria
[e] FEMTO-ST, MN2S, UTBM, Site de Sévenans, 90010 Belfort Cedex, France

ARTICLE INFO

Article history:
Received 13 May 2018
Received in revised form
7 August 2018
Accepted 10 August 2018
Available online 13 September 2018

Keywords:
Intermetallic compounds
Datamining approach
Principal component analysis
Artificial neural network
Formability

ABSTRACT

In this paper, two techniques of datamining tools were adopted, a principal component analysis (PCA) and artificial neural network (ANN). A PCA to classify, select and identify several combinations between transition element A and B (B = Ti, Zr, Hf, Sc, Y, La and Th) and ANN to predict ΔH for ternary hydrides. Based on the datasets selected from different works, a principal component analysis (PCA) has been applied to select, classify and identify around 76 possible combinations between transition metal elements A and B. The results showed that the clustering of combinations A-B are significantly influenced by the atomic parameters of element A, such atomic radius (R_A), Pauling's electronegativity (χ_A) and atomic electron density (Z_A/R_A^3). From 76 combinations, 55 systems which have $\chi_A \geq 1.5$, $Z_A/R_A^3 > 1.28$ and $R_A < 1.46$ Å are categorized as group 1. On the other hand, 21 systems which have $\chi_A < 1.5$, $Z_A/R_A^3 < 1.28$, and $R_A > 1.46$ Å are categorized as group 2. From the first group, 46 different combinations are identified and have a negative ΔH, within 18 well-known promising binary alloys of hydrogen storage.

An (6-15-1) architecture of artificial neural network (ANN) has been developed to estimate the ΔH for the other ternary hydrides selected from different published works. The performance indices such as relative error, coefficient of determination (R^2) and mean square error (MSE) were used to control the performance of obtained results. In addition to this, the ΔH obtained from ANN model were compared with those experimental data and theoretical results available in the literature.

* *Corresponding author*. Unité de Recherche Appliquée en Energies Renouvelables, URAER, Centre de Développement des Energies Renouvelables, CDER, 47133, Ghardaïa, Algeria.
E-mail address: benyelloul_kamel@yahoo.fr (K. Benyelloul).
https://doi.org/10.1016/j.ijhydene.2018.08.122

Introduction

The intermetallic compound hydrides are considered as an attractive material for hydrogen storage. These hydrides are promising materials both as fuel as in rechargeable batteries, fuel cells and heat storage [1,3]. They have been received wide attention due to their large number of technical applications and they opened new horizons for industrial development due to their widespread in the development of hydrogen-absorbing metal alloys [4–6]. In particular, the intermetallic compound hydrides (AB_5H_x, ABH_x, AB_2H_x, A_2BH_x and AB_3H_x, as well as solid solution alloys) are also widely investigated and have been an active research area due to their excellent hydrogen absorbing properties [7,8]. Indeed, this hydride is based on the combination between transition elements A and B. The stability of the binary formation A-B depends on both the atomic parameters of transition elements (A, B) and the processing conditions (Pressure, temperature,…). Due to the complex and non-linear effect of parameters influencing the stability and thermodynamic properties in the formation of both intermetallic compounds and their hydrides, a large number of theoretical studies and experimental works have been made [9–12], to understand, classify and study of various aspects of intermetallic compounds forming the hydrides.

The alloying ability of intermetallic compounds can be evaluated by its enthalpy of formation (ΔH). A negative ΔH means an exothermic process, and the intermetallic compounds can be formed and stable. For the predicting stable intermetallic compounds and its hydrides several theoretical methods have been reported to estimate the ΔH, such the Miedema's model [10,13], a semi-empirical model [11,14] and density functional theory (DFT) calculation [9]. Herbst [10] has used Miedema's model to estimate ΔH for binary $AH_{x\ (x=1\text{-}3)}$ and $AB_nH_{x\ (n=1,2,3,5)}$ type hydrides and predict the hydrogen content of binary and ternary hydrides. Griessen et *al.* [11] have calculated by means of a semi-empirical model based on the electronic structure of the host alloys the heats of formation of the transition metal hydrides A_5BH_x, A_2BH_x, ABH_x, AB_2H_x and AB_5H_x, where A = Sc, Ti, V, Y, Zr, Nb, La, Hf, Ta, Ni, Pd and B is transition metal, and have also predicted 44 different compounds to react with hydrogen. Watson and Bennett [15] have predicted the ΔH for 276 equiatomic composition of transition metal alloys by an electron band theory model based on the bandwidth, Fermi level position and number of electrons in the band. Also, by mean of the DFT calculation, J.Wang et *al* [9] have estimated the enthalpies of formation for Al-X (X = Co, Cu, Hf, Mg, Mn, Ni, Sr, V, Ti, Y and Zr) systems. However, the abundance of physical parameters of two transition elements A and B are probably responsible for the difficulties in the study of the formability of A-B system. It is difficult to develop them in the form of mathematical equations. Van Mal [16] proposed some criteria for classification of ternary transition metal hydrides according to ΔH.

Therefore, it is important to understand the effect of atomic parameters on the formability of intermetallic compounds and its hydrides. In this context, much recent informatics methods have been developed. In which a new knowledge system is built by collecting and classifying information with the help of calculations and databases [2]. Among this informatics method which attracts an increasing care of the material science researches is the data mining (DM). DM approach becomes a significant technique in the building knowledge-based system by inductive inference [17]. The most popular DM approaches are partial least square (PLS) (Linear regression), artificial neural network (ANN) and principal component analysis (PCA).

Following specific interest, a recent example of this type of approaches comes from Benyelloul et *al.* [18–20] where the authors demonstrate an application of ANN and PCA for the task of predicting several physical parameters and thermodynamic properties. In other hand, Jin Guo et *al* [21] have used the partial least square (PLS) method to classify and select the binary transition metal alloys. They found that the stability and the formability of hydrides of binary transition metal alloys are significantly influenced by electron density.

However, it is shown that the atomic parameters (such as electron density, atomic radius, Pauling's electronegativity and molar volume) effect strongly on the formation of binary alloys A-B systems and it is mandatory to study this effect.

Hence, two distinct types of datamining methods have been proposed in this study. Firstly, a PCA is used to select and classify several combinations between transition element A and B favourable for stability. Secondly, an ANN is used to predict ΔH from several ternary hydrides.

In this view, our paper is structured as follows. In section Computational methods, we give a brief overview of basic techniques used in datamining. In the following section concerning the results and discussion, we analysed the obtained results. Finally, a conclusion of the present work is given.

Computational methods

Principal component analysis

Principal component analysis (PCA) is one of most technique in exploratory data analysis in DM approach, it is the technique for data study which involves many variables and can effectively solve the correlation problem [22,23]. PCA is used to reduce a large set of variable that still contains most of the information in the large data set [24]. The original matrix data is decomposed and projected into two plots, the sample are classified in the score plot and the descriptors in terms of their separation of the samples in the loading plot following the principle component (PCs) axis [18]. The details of the PCA can be summarized as follow steps [23,25].

Step 1: The original multiple quality characteristic matrix:

$$X^0 = \begin{bmatrix} X_{11} & . & . & X_{1n} \\ . & . & . & . \\ . & . & . & . \\ X_{m1} & . & . & X_{mn} \end{bmatrix} \quad (1)$$

where m is the number of trials and n is the number of the quality characteristic.

Step 2: Data standardization

$$\tilde{X}_{ij} = \frac{X_{ij} - \tilde{X}_j}{S_j} \quad (2)$$

in which $\tilde{X}_j$ is the mean value of X_j and S_j is the standard deviation of X_j. Thus the standard matrix is:

$$X = \begin{bmatrix} \tilde{X}_{11} & . & . & \tilde{X}_{1n} \\ . & . & . & . \\ . & . & . & . \\ \tilde{X}_{m1} & . & . & \tilde{X}_{mn} \end{bmatrix} \quad (3)$$

Step 3: Correlation coefficient matrix

$$R = \begin{bmatrix} r_{11} & . & . & r_{1n} \\ . & . & . & . \\ . & . & . & . \\ r_{m1} & . & . & r_{mn} \end{bmatrix} \quad (4)$$

$$\text{where } r_{jk} = \frac{\sum_{i=1}^{n}(X_{ij} - \tilde{X}_j)(X_{ik} - \tilde{X}_k)}{\sqrt{\sum_{i=1}^{n}(X_{ij} - \tilde{X}_j)^2 \sum_{i=1}^{n}(X_{ik} - \tilde{X}_k)^2}} \quad (5)$$

in which $\tilde{X}_j$ and $\tilde{X}_k$ are the mean values of X_j and X_k, respectively.

Step 4: The determination of eigenvalues and eigenvectors are in following equation

$$(R - \lambda_j)V_j = 0 \quad (6)$$

where, λ_j is the eigenvalues, $\sum_{j=1}^{n}\lambda_j = n$, $j = 1, 2, \ldots, m$; and $V_j = (a_{j1}, a_{j2}, \ldots, a_{jn})$ is the eigenvectors corresponding to the eigenvalues.

Step 5: Explained variations.

$$b_j = \frac{\lambda_j}{\sum_{j=1}^{n}\lambda_j} \text{ and } a_p = \frac{\sum_{j=1}^{p}\lambda_j}{\sum_{j=1}^{n}\lambda_j} \quad (7)$$

in which, b_j represents the contribution rate (CR) of the j_{th} eigenvalue, and a_p represents the cumulative contribution rate (CCR) of the p_{th} eigenvalue. Eigenvectors correspond to the eigenvalues.

Step 6: Final data.

The final data is formulated as:

$$Y_{mj} = \sum_{i=1}^{n} X_m(i)V_{ij} \quad (8)$$

in which, Y_{mj} is the j_{th} final data formulated from eigenvector V_{ij} of principal components in the analysis.

Finally, based on its contribution rate and cumulative contribution rate to the total variance, the principal components (PCs) are ordered. The first principal component has the largest possible variance and each succeeding PC are determined with the property that they are orthogonal to the first PC and that the highest variance possible [23].

Artificial neural network

An artificial neural network (ANN) is a biologically inspired processing unit that is extensively used for many areas to identify and to model nonlinear system [26]. The ANN plays the role on the relationship between input *x* and output *y*. The typical ANN architecture is composed by layers, the first takes input from outside the ANN, and the last outputs the final results to the user, while layers in between are called hidden and communicate only with other layers [27].

The relationship between the inputs and output is given by Refs. [19,28]:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + \theta\right) \quad (9)$$

in which, x_i and y are the inputs and output respectively, w_i the weight of the neural model, θ is the bias and f is the activation function.

The main of training process is to found values of weights of the neurons (w_i), which minimize an error between y and predicting y (y_{pred}).

In this paper, we have used a multilayer perceptron MLP [29] as ANN architecture, to predict the ΔH. More details about the artificial neural network method can be found in our previous paper [18].

In order, to evaluate the performance of training process different error indexes were used, a mean square error (MSE) and coefficient of determination (R^2), the computational equations of these parameters are given by:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - y_{i(pred)}\right)^2 \quad (10)$$

$$R^2 = 1 - \left(\frac{\sum_{i=1}^{N}\left(y_i - y_{i(pred)}\right)^2}{\sum_{i=1}^{N}\left(y_{i(pred)}\right)^2}\right) \quad (11)$$

where N is the number of training patterns, y_i is the desired output and $y_{i(pred)}$ is the predicted network output.

Results and discussion

Principal component analysis (PCA) results

In order to select, classify and identify the intermetallic compounds favourable for stability of hydrides, a PCA has been performed for the database including ~76 combinations between transition metal elements A and B and their corresponding atomic parameters (descriptors), such as the atomic radius R_A, R_B, the Pauling's electronegativity χ_A, χ_B, and the atomic electron density Z_A/R_A^3, Z_B/R_B^3, respectively. ($Z_{A,B}$ is the number of valence electrons for element A and B, respectively). The values of data were collected from reference [21,30].

The contribution rate (CR) and cumulative contribution rate (CCR) among all principal components (PCs) are given in Fig. 1. It is indicated that the first two PCs cover ~82% of the total variance. The first principal component (PC1) explained 44%, of the variance in the datasets, whereas the second principal component (PC2) captures 38%.

The obtained plots of first two PCs from PCA of a set of data which contain the combinations (A-B) between transition metal elements A and B are shown in Fig. 2 (a,b). The properties with similar PC value are highly correlated, while inverse PC values indicate inverse correlation.

The correlation between descriptors can be evaluated through the loading plot which is depicted in Fig. 2(a). The effect of each parameter was examined. Whereas, strong correlation between Pauling's electronegativity χ_A, χ_B and atomic electron density Z_A/R_A^3, Z_B/R_B^3, can be seen respectively. It can also be observed that PC2 has captured the variation in R_A, χ_B, Z_B/R_B^3 and R_B, χ_A, Z_A/R_A^3. We note that the value of

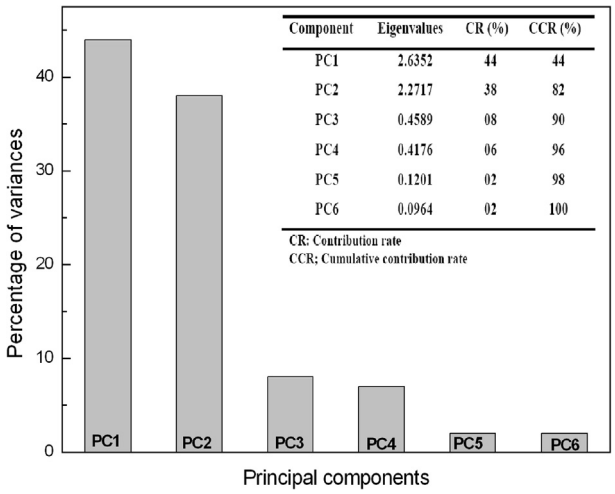


Fig. 1 – Total variance of PCA results.

atomic radius (R_A, R_B) is inversely correlated to the Pauling's electronegativfity (χ_A, χ_B), respectively. Accordingly, it confirms that element which have large (small) atomic radius have low (high) Pauling's electronegativity.

The score plot for the first two principal components (PC1 and PC2) can be interpreted by the Fig. 2(b). It can be observed that the separation of the two groups is rather clear-cut (diagonal separation). Hence, the whole of the combinations (A-B) is classified into two groups. (The points and hollow circle corresponds to the combinations A-B). Following the mentioned interpretations, it can be remarked that the separated groups of the combinations (A-B) are strongly influenced by the parameters, such as atomic radius (R_A, R_B), Pauling's electronegativity (χ_A, χ_B) and atomic electron density (Z_A/R^3_A, Z_B/R^3_B). The combinations, with large atomic radius R_A, low Pauling's electronegativity χ_A and low atomic electron density Z_A/R^3_A from element A, small atomic radius R_B, high Pauling's electronegativity χ_B and high atomic electron density Z_B/R^3_B from atom B, have a positive PC2. While the combinations A-B with small atomic radius R_A, high Pauling's electronegativity χ_A and high atomic electron density Z_A/R^3_A, from element A, large atomic radius R_B, low Pauling's electronegativity χ_B and low atomic electron density Z_B/R^3_B from atom B, where, they have negative PC2.

A high correlation between the clustering and the atomic parameters of element A has been observed in the previous analysis. Hence, in order to identify the location of the groups with respect to the second principal component axis (PC2), our current study is focusing on the variation between the combinations A-B, atomic radius (R_A), Pauling's electronegativity (χ_A) and atomic electron density (Z_A/R^3_A). Accordingly, a PCA projection score plot for element A and their atomic parameters (χ_A, Z_A/R^3_A and R_A) in the several A-B combinations are depicted in Fig. 3(a,b,c,d), respectively. Also, for the better illustration the clusters are depicted separately in Fig. 2 (c, d).

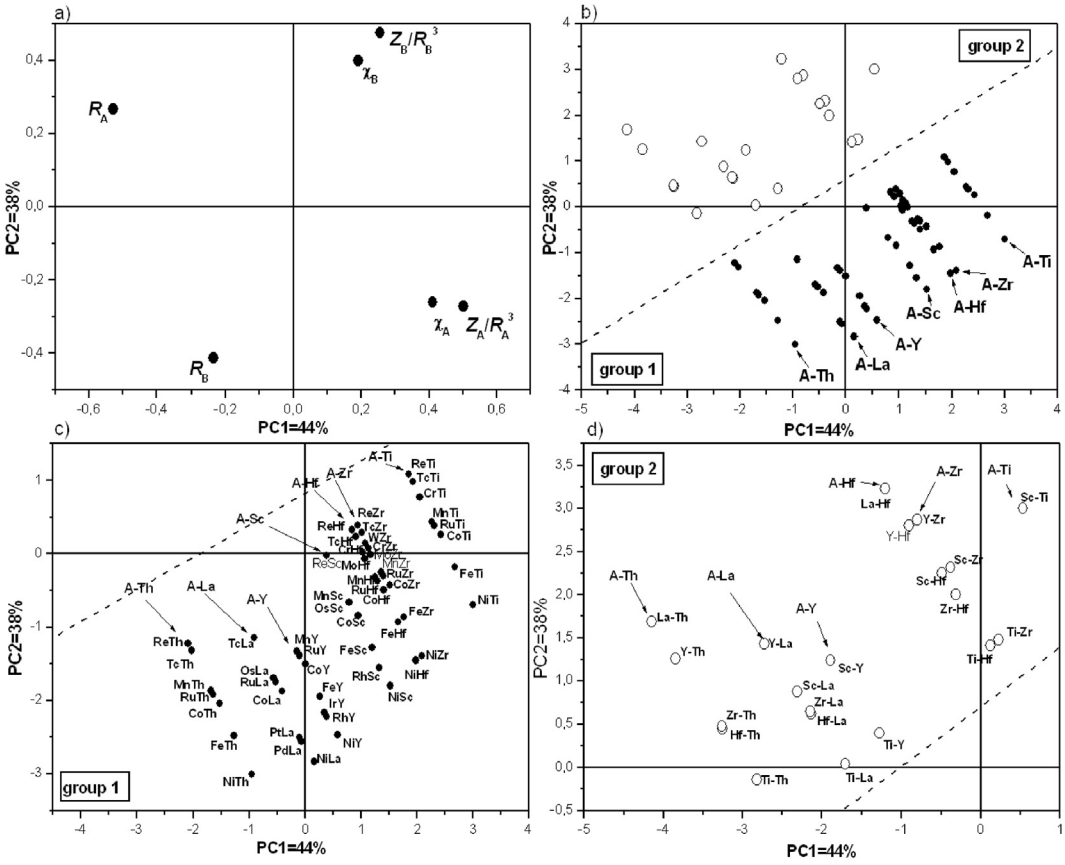


Fig. 2 – PCA results a) PCA loading plot, b) PCA score plot for 76 different combinations, c) PCA score plot for 55 combinations, d) PCA score plot for 21 combinations.

Throughout the Fig. 2 (c), and Fig. 3(a,b,c,d), we can observe that the element A for different combinations A-B, have the atomic parameters χ_A, Z_A/R_A^3 and R_A values, greater than or equal 1.5, greater than 1.28 and less than 1.46 Å [21,30], respectively, (except for Mn–B (where B = Ti, Zr, Hf, Sc, Y and Th), Mn have χ_A equal to 1.5). Meanwhile, from Fig. 2 (d) and Fig 3(a,b,c,d), the Sc–Ti, A-Zr (where A = Ti, Sc, Y), A-Hf (where A = Ti, Zr, Sc, Y, La), A-Y (where A = Ti, Sc), A-La (where A = Ti, Hf, Zr, Sc, Y) and A-Th (where A = Ti, Hf, Zr, Y, La) systems, the values of χ_A, Z_A/R_A^3 and R_A are less than 1.5, less than 1.28 and greater than 1.46 Å, respectively. Based on the above interpretation, we can say that the systems which have $\chi_A \geq 1.5$, $Z_A/R_A^3 > 1.28$ and $R_A < 1.46$ Å are categorized as group 1. On the other hand, the systems which have $\chi_A < 1.5$, $Z_A/R_A^3 < 1.28$, and $R_A > 1.46$ Å are categorized as group 2. Effectively, according Fig. 2 (b, c and d), we can observe that out of 76 combinations, 55 combinations are categorized as group 1 and 21 combinations are categorized as group 2.

The ΔH, which is the most thermodynamic parameter, used to predict the stability of the combination systems. For $\Delta H < 0$ kJ/mol H_2 the system is stable and for $\Delta H > 0$ kJ/mol H_2 the system is unstable. In order to select the stable and unstable systems for our studied combinations, the obtained classification are compared to the experimental data and the theoretical value of ΔH for intermetallic compounds reported by Watson et *al* [15], Griessen et *al* [11] and Van Mal [16]. On 55 combinations from group 1, 46 were observed that have negative ΔH, indicating that the mentioned combinations are stable relative to the constituent elements and constitute an intermetallic compound. Accordingly, the stable and their corresponding intermetallic compounds are displayed in Table 1. Also, from this Table, 8 systems have classified as the second group have a positive or no ΔH, indicating that the combinations are unstable. For the remaining combinations, no experimental or theoretical ΔH for intermetallic compounds are available in literature for comparison.

In other hand, it can be clearly shown from Fig. 2(c), that well-known promising binary alloys of hydrogen storage such Fe–Ti, Ni–Ti, Co–Ti, Cr–Ti, Fe–Sc, Ni–Zr, Cr–Zr, Ni–La, Co–La, Ni–Y, Co–Y, Fe–Y, Ni–Th and Co–Th [7,10,16,31] appear in the cluster 1(see Table 1). Fe–Ti, Ni–Ti and Co–Ti with cubic structure (CsCl-type structure), Ni–Zr, Fe–Y and Fe–Sc with orthorhombic *Cmcm* space group, cubic C15 (Fd-3m) and hexagonal C14 ($P6_3/mmc$) [31,34], respectively, are capable of absorbing hydrogen reversibly. These type of alloys will react directly with hydrogen to form a $FeTiH_{x\ (x=1\ and\ 1.6)}$, $NiTiH_{x\ (x=1.0\ and\ 1.4)}$, $CoTiH_{1.5}$, $NiZrH_{x\ (x=2\ and\ 3)}$, $Fe_2YH_{x(x=3,4\ and\ 5)}$ and Fe_2ScH 2 hydrides, respectively [10,31,35]. Also, the binary alloys NiLa and CoLa with Cu_5Ca type structure [7,8] have been attracted the care, in the last years, regarding their high hydrogen absorption/desorption capacity, good cyclic abilities, fast kinetics and low equilibrium pressures [7]. The most investigated of these type intermetallic compounds are Ni_5La and Co_5La and these produce

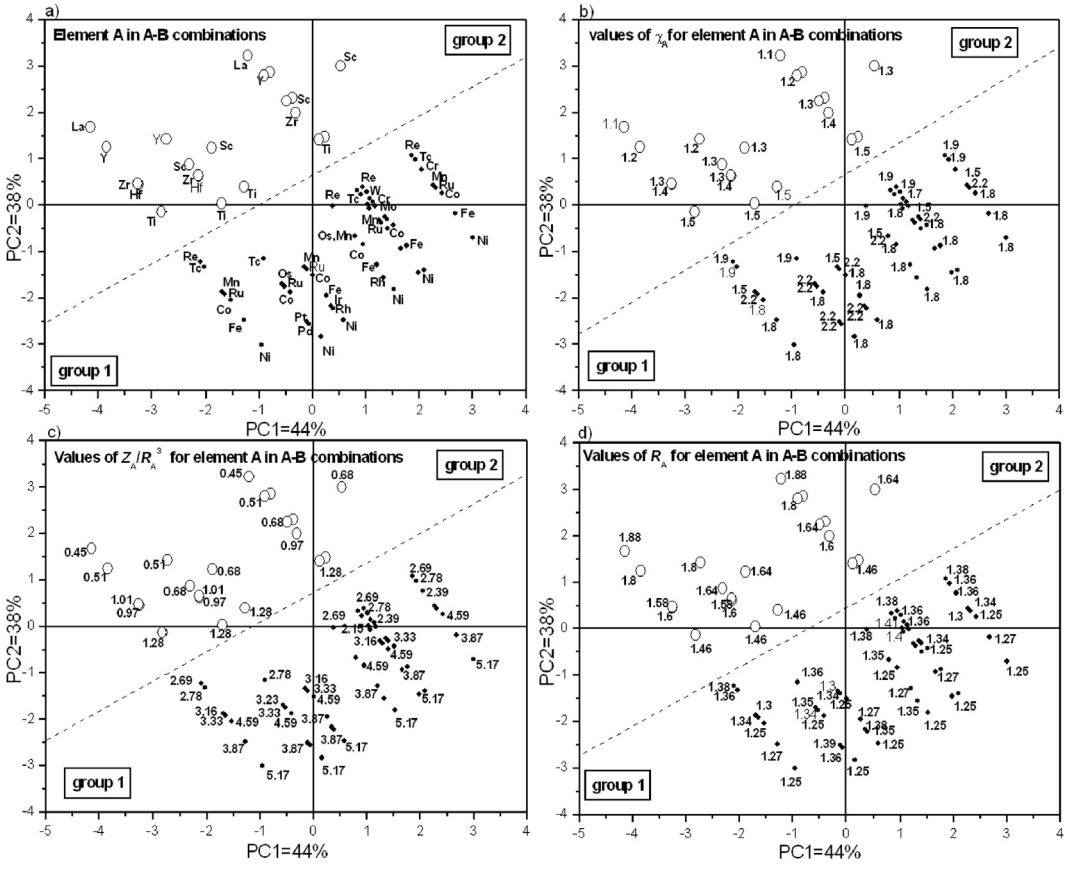


Fig. 3 – PCA score plot for atomic parameters of element A in A-B combinations a) the elements A in A-B systems, b) the Pauling's electronegativity (χ_A) value, c) the atomic electron density (Z_A/R_A^3) value, d) the atomic radius R_A values.

Table 1 – Group 1 and group 2 for different combinations (A-B).

Combinations	Group 1 ($\Delta H < 0$)	Inter. comp	Group 2 ($\Delta H \geq 0$)	Combinations	Group 1 ($\Delta H < 0$)	Inter. comp	Group 2 ($\Delta H \geq 0$)
A-Ti	NiTi[a]	NiTi[b]		A-Sc	NiSc[a]	Ni_2Sc[b]	
	FeTi[a]	FeTi[b]			RhSc[a]		
	CoTi[a]	CoTi[b,c]			FeSc[a]	Fe_2Sc[b,f]	
	RuTi[a]				CoSc[a]	Co_2Sc[b]	
	MnTi[a]	Mn_2Ti[b]			OsSc[a]		
	CrTi[a]	Cr_2Ti[b]			ReSc[a]	Re_2Sc[b]	
	TcTi[a]				MnSc[a]		
	ReTi[a]	ReTi[b]		A-Y	NiY[a]	Ni_5Y[b]	
			ScTi[a]		RhY[a]		
A-Zr	NiZr[a]	NiZr[d,f]			IrY[a]		
	FeZr[a]	Fe_2Zr[b]			FeY[a]	Fe_2Y[f]	
	CoZr[a]	CoZr[c]			CoY[a]	Co_5Y[b]	
	RuZr[a]				RuY[a]	Ru_2Y[b]	
	MnZr[a]	Mn_2Zr[d]			MnY[a]		
	MoZr[a]						ScY[a]
	CrZr[a]						TiY[a]
	WZr[a]	W_2Zr[b]		A-La	NiLa[a]	Ni_5La[b,f]	
	TcZr[a]				CoLa[a]	Co_5La[b]	
	ReZr[a]	Re_2Zr[b]			RuLa[a]	RuLa[b]	
			ScZr[a]			Pt_5La[d]	
			TiZr[a]	A-Th	NiTh[e]	NiTh[e]	
			YZr[a]		CoTh[e]	Co_5Th[e]	
A-Hf	NiHf[a]	NiHf[b]				Fe_3Th_7[e]	
	FeHf[a]						
	CoHf[a]	CoHf[b,c]					
	CrHf[a]	Cr_2Hf[b]					
	RuHf[a]						
	MnHf[a]	Mn_2Hf[b]					
	MoHf[a]	Mo_2Hf[b]					
	TcHf[a]						
	ReHf[a]						
			YHf[a]				
			TiHf[a]				

[a] Ref [15].
[b] Ref [10].
[c] Ref [32].
[d] Ref [33].
[e] Ref [16].
[f] Ref [31].

$Ni_5LaH_{6.7}$ and Co_5LaH_9 [10] hydrides, respectively. For Ni–Y, Co–Y, Cr–Ti, Cr–Zr, Fe–Y, Ni–Th and Co–Th combinations their hydrides are respectively Ni_3YH_4, Co_3YH_4, $Cr_2TiH_{1.2}$, $Cr_2ZrH_{3.8}$, $Fe_{23}Y_6H_{21.5}$ [10], $Ni_2ThH_{2.3}$ [36] and $CoThH_{4.2}$ [37]. In addition to that, the binary alloys Pd–La, Pt–La, Fe–Th and Mn–Th can form the stable hydrides, such Pd_5LaH_6 [16], $Pt_5LaH_{1.2}$ [10], $Fe_3ThH_{1.7}$ [16] and $Mn_{23}Th_5H_{24}$ [36], respectively. The 18 well-known hydrides are projected from PCA score plot, and illustrated in Fig. 4.

In addition, from Fig. 2 (c) we perceive that 5 systems A-La (A = Os, Tc) and A-Th (A = Ru, Tc, Re), are categorizing as group 1. Unfortunately, to our knowledge, there are no experimental or theoretical data available. Based on the above PCA analysis, it appears very probable that the Os–La, Tc–La, Ru–Th, Tc–Th and Re–Th systems are favourable for stability and can form the corresponding hydrides from their type alloys. The as-obtained combinations can be considered as the predicted intermetallic compounds hydrides and are a

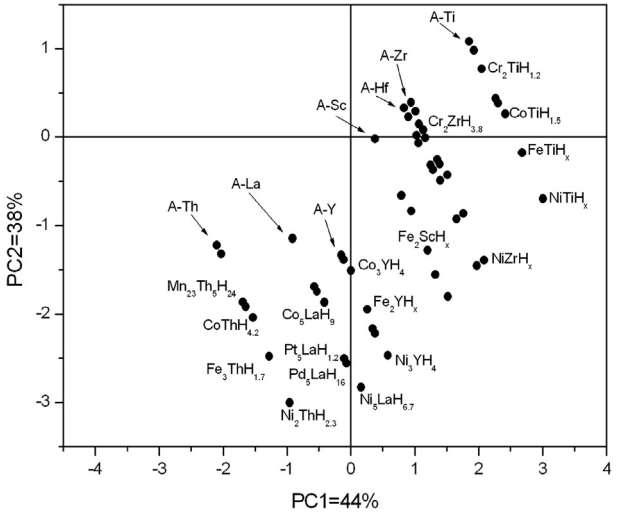


Fig. 4 – Illustration for 18 well-know hydrides in PCA score plot.

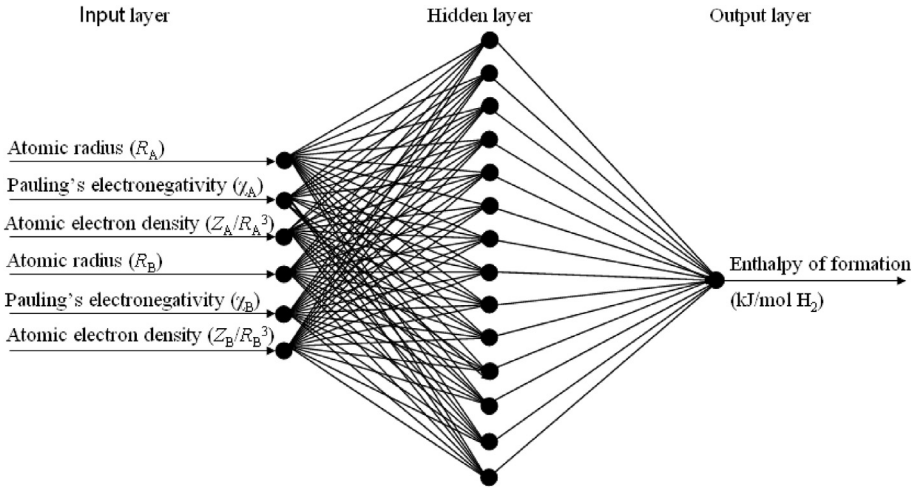


Fig. 5 – The ANN (6-15-1) model for predicting the enthalpy of formation.

reference for future experimental and/or theoretical research works.

The combined effect of atomic parameters (χ_A, Z_A/R_A^3 and R_A) of elements A for the ~76 mentioned combination systems A-B on the stability and the hydrogen storage properties is interpreted. The result has shown two groups of combinations into formed and unformed systems. Whose ~72% for datasets categorizing in first group, corresponding to the 55 combinations A-B which have an atomic parameters of transition element A, $\chi_A \geq 1.5$, $Z_A/R_A^3 > 1.28$ and $R_A < 1.46$ Å. From those 55 systems (72%), 46 combinations were found to have a negative ΔH and can form an intermetallic compounds, within 18 well-known promising binary alloys of hydrogen storage.

For previous results by the classification method, the stability and the formability of intermetallic compounds hydrides is strongly influenced by the atomic radius, Pauling's electronegativity and atomic electron density. Accordingly, on the below section we have developed a predictive method for the estimation of ΔH for ternary hydrides by applying artificial neural network (ANN).

Artificial neural network (ANN) results

In this section, a predictive method based on artificial neural network (ANN) has been used to estimate the ΔH of ternary hydrides. We collected data from available literature [10,38,39] and constructed property database. Database with

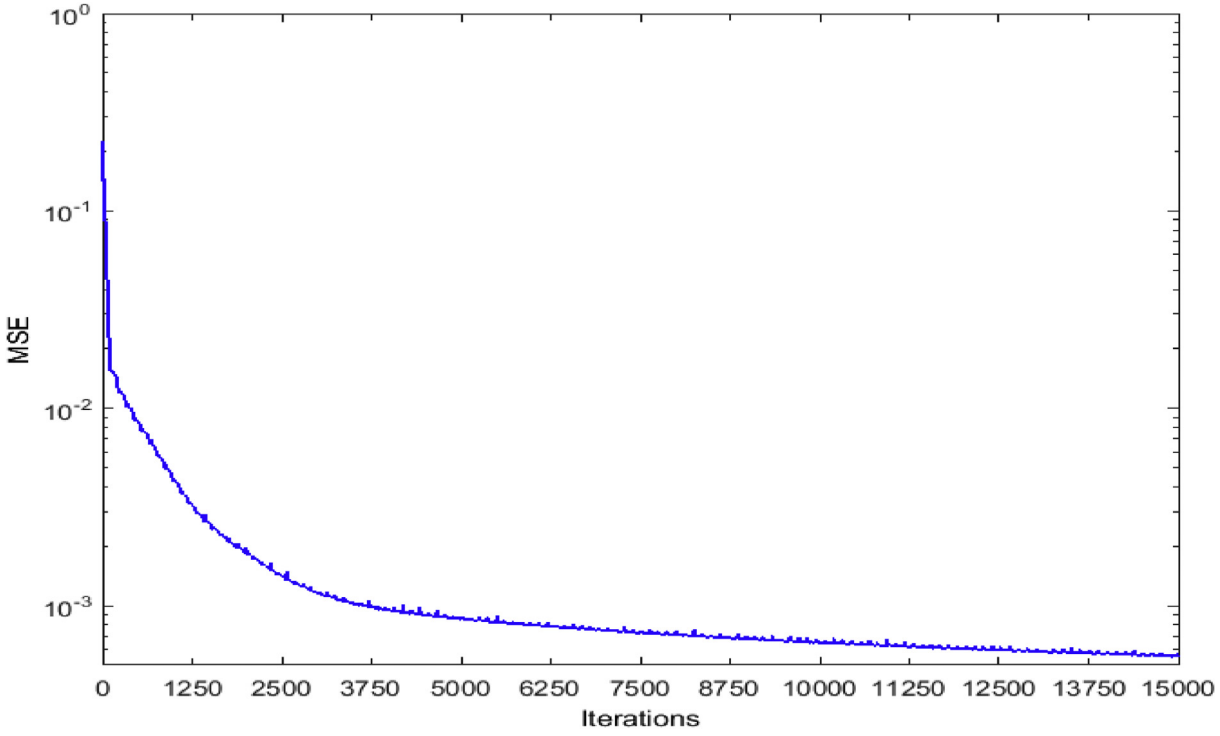


Fig. 6 – Training artificial neural network model for the enthalpy of formation.

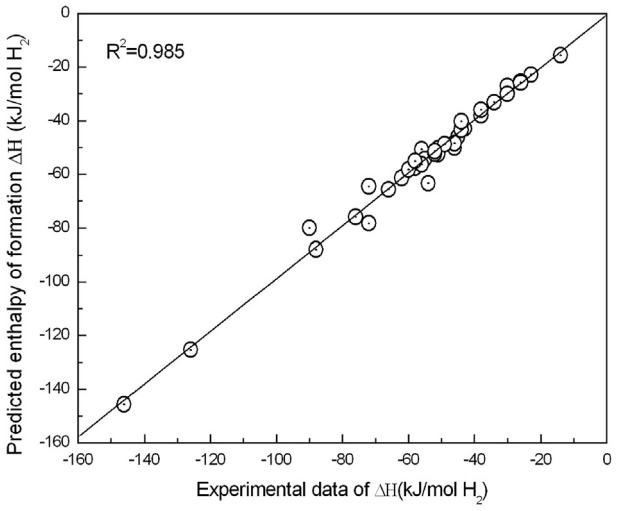


Fig. 7 – ANN prediction versus experimental data for the enthalpy of formation.

39 ternary hydrides is interpreted, including the atomic radius (R_A, R_B), the Pauling's electronegativity (χ_A, χ_B) and the atomic electron density (Z_A/R_A^3, Z_B/R_B^3) element A and B forming these ternary hydrides and the ΔH of the corresponding hydrides. R_A, R_B, χ_A, χ_B, Z_A/R_A^3, Z_B/R_B^3 as inputs nodes and ΔH as output node were selected for the MLP network. For the training process, all inputs and outputs in the datasets are normalized in the interval [−0.5, +0.5], according to the following relationship:

$$x_{normalized} = \frac{(x - x_{min})}{(x_{max} - x_{min})} - 0.5 \quad (12)$$

where x is the original value from database, x_{min} and x_{max} are the minimum and the maximum value within the entire value data for x, $x_{normalized}$ is the normalized value.

In order to assign an appropriate number of hidden layers, nodes to the hidden layer and the transfer function, many training processes were carry out. The suitable number of hidden layers is determined by a trial and error method, mean square error (MSE) was chosen in error calculation. The selected MLP architecture that provides good training and testing is composed of one hidden layer with 15 nodes. Therefore, an ANN with architecture 6-15-1 (6 inputs, 15 hidden neurons and one output) was used to estimate ΔH. The transfer function between the input layer and hidden layer is tangent sigmoid transfer function $(\tan sig(x) = 2/(1 + \exp(-2x)) - 1)$, while the activation function between the hidden layer and the output is linear transfer function. (*purelin*). The architecture of ANN is illustrated in Fig. 5.

After training and testing processes, the performance of the ANN is analysed. The training process of artificial neural network is illustrated in Fig. 6. From this figure, we can notice that the average mean square error (MSE) decreased with increase the number of iterations. The curve indicates that the convergence with MSE equal to 5.651×10^{-4} is achieved after 15000 training epochs. The performance of ANN model is evaluated and illustrated in Fig. 7, it can be seen that the R^2 is 0.985 (close to 1), indicating good agreement between datasets values and predicted values of ΔH for the ternary hydrides.

The estimated values of ΔH as well as the corresponding experimental and theoretical value are displayed in Table 2. The relative error in Table 2 is defined as the absolute difference between experimental and predicted ΔH divided by

Table 2 – Comparison of the predicted and experimental ΔH (kJ/mol H_2) for the ternary hydrides.

system	Exp.	Other	ANN	Rel. err.	System	Exp.	Other	ANN	Rel. err
$Ni_3CaH_{4.5}$[a]	−58[b]	−62[a]	−57	0.017	$Ni_2CeH_{3.9}$[a]	−72[b]	−62[a]	−78	0.083
$Ni_5CaH_{4.5}$[a]	−38[b]	−50[a]	−38	0	$Ni_3CeH_{4.2}$[a]	−44[b]	−32[a]	−40	0.091
$Fe_2ScH_{3.2}$[a]	−30[b]	−89[a]	−29	0.033	Ni_5CeH_6[a]	−14[b]	−15[a]	−16	0.143
$Cr_2TiH_{3.2}$[a]	−23[b]	−78[a]	−22	0.043	Cu_5NdH_3[a]		−51[a]	−51	0
$FeTiH_{1.6}$[a]	−34[b]	−86[a]	−33	0.029	$Ni_5Eu\ H_{5.5}$[a]	−26[b]	−51[a]	−25	0.038
$CoTiH_{1.5}$[a]	−62[b]	−71[a]	−61	0.016	Gd_2MnH_3[a]	−88[c]	−138[a]	−88	0
$CuTiH$[a]	−76[b]	−114[a]	−75	0.013	$GdFe_3H_{3.2}$[a]	−51[c]	−92[a]	−52	0.020
Co_3YH_4[a]	−56[c]	−47[a]	−51	0.089	$Co_2GdH_{4.5}$[a]	−54[b]	−63[a]	−63	0.167
Ni_3YH_4[a]	−46[b]	−28[a]	−51	0.109	Co_3GdH_5[a]	−52[c]	−42[a]	−52	0
$Mn_2ZrH_{3.6}$[a]	−44[b]	−103[a]	−43	0.023	$Co_5GdH_{2.2}$[a]	−30[b]	?1[a]	−30	0
$NiZrH_{2.8}$[a]	−66[b]	−78[a]	−66	0	$Ni_2GdH_{4.1}$[a]	−90[c]	−56[a]	−80	0.111
Co_5LaH_9[a]	−45[c]	−18[a]	−46	0.022	$Ni_5GdH_{2.9}$[a]	−26[b]	+12[a]	−26	0
$NiLaH_4$[a]	−126[b]	−87[c]	−125	0.008	$GdRu_2H_{3.7}$[a]	−60[c]	−51[a]	−58	0.033
$Ni_2LaH_{4.6}$[a]	−55[c]	−56[a]	−54	0.018	$Gd_2RhH_{3.3}$[a]	−49[c]	−33[a]	−49	0
$Ni_5LaH_{6.7}$[a]	−30[c]	−16[a]	−27	0.100	$DyRu_2H_{3.1}$[a]	−56[b]	−54[a]	−56	0
$Cu_5LaH_{3.72}$[a]	−43[a]	−59[a]	−43	0	Fe_2ErH_4[a]	−58[c]	−88[a]	−55	0.052
$Rh_2LaH_{4.9}$[a]	−44[c]	−26[a]	−43	0.023	$NiYbH_{2.7}$[a]	−146[b]	−126[a]	−145	0.007
$Co_2CeH_{4.1}$[a]	−72[b]	−74[a]	−64	0.111	$Ni_2YbH_{3.1}$[a]	−52[b]	−93[a]	−51	0.019
$Co_3CeH_{3.8}$[a]	−46[b]	−51[a]	−48	0.043	Ni_5YbH_3[a]	−26[b]	−48[c]	−26	0
Co_5CeH_3[a]	−38[b]	−22[a]	−36	0.053					

[a] Ref [10].
[b] Ref [39].
[c] Ref [38].

the experimental value. Except for Cu_5NdH_3, the relative error was calculated between predicted value and theoretical value obtained by Herbst [10]. From this Table, the median relative error of our predictions is 0.0388, which means half of the predictions have a relative error that is less than 3.88%. There are seven ternary hydrides whose prediction relative error is better to be 08%, namely 82% of the predictions have a relative error that less than 08%. Only two ternary hydrides ΔH are predicted with a relative error slightly above 14%. These hydrides are $Co_2GdH_{4.5}$ and Ni_5CeH_6. It is seen that the obtained ΔH are basically consistent with the experimental data [38,39], indicating that the constructed ANN (6-15-1) model is capable to predict accuracy ΔH as a function of the atomic radius, Pauling's electronegativity and atomic electron density.

Conclusion

Through this work, two techniques of datamining tools were adopted, a principal component analysis (PCA) and artificial neural network (ANN). A PCA to classify, select and identify ~76 possible combinations between transition element A and B (B=Ti, Zr, Sc, Hf, Y, La and Th) and ANN to predict ΔH for ternary hydrides. Therefore, from the obtained selection and classification results for PCA analysis, it has been shown that the different combinations (A-B) are significantly influenced by the atomic radius (R_A), Pauling's electronegativity (χ_A) and atomic electron density (Z_A/R_A^3) for element A. The separation of the two groups showed a light clear-cut which consist of combined transition metal element A and B. Among the 76 different combinations, 55 (72%) are summarize as group 1 and have $\chi_A \geq 1.5$, $Z_A/R_A^3 > 1.28$ and $R_A < 1.46$ Å. Otherwise, a combination systems (A-B) categorizing as group 2. Within 55 systems, 46 different combinations are identified and have a negative ΔH, among them the well-known promising intermetallic compounds of hydrogen storage alloys.

According to the obtained results of the Os–La, Tc–La, Ru–Th, Tc–Th, and Re–Th systems that classified as cluster 1, it can be inferred from PCA analysis that the Os–La, Tc–La, Ru–Th, Tc–Th, and Re–Th systems can be formed and constructed the corresponding hydrides, according to their type alloys. Furthermore, a predictive method based on artificial neural network (ANN) has been used to estimate ΔH of ternary hydrides. Through the low relative errors obtained between experimental data and predicted ΔH and best value of coefficient of determination R^2 (0.985), the developed ANN model showed remarkable results.

Accordingly, the good validated results, obtained by the PCA analysis of the selection, classification and identification of the intermetallic compounds favourable for stability of hydrides, as well as the accuracy satisfactory results were confirmed by the ANN model prediction of ΔH. Consequently, these results seem to be powerful potential in applying data-mining techniques in the materials science.

Annex

Database of combinations A-B calculations Ref [21,30].

A	B	R_A	X_A	Z_A/R_A^3	R_B	X_B	Z_B/R_B^3
Cr	Ti	1,36	1,6	2,39	1,46	1,5	1,28
Cr	Zr	1,36	1,6	2,39	1,6	1,4	0,97
Cr	Hf	1,36	1,6	2,39	1,58	1,3	1,01
Co	Y	1,25	1,8	4,59	1,8	1,2	0,51
Co	La	1,25	1,8	4,59	1,88	1,1	0,45
Co	Ti	1,25	1,8	4,59	1,46	1,5	1,28
Co	Zr	1,25	1,8	4,59	1,6	1,4	0,97
Co	Hf	1,25	1,8	4,59	1,58	1,3	1,01
Co	Sc	1,25	1,8	4,59	1,64	1,3	0,68
Co	Th	1,25	1,8	4,59	1,8	1,1	0,52
Fe	Ti	1,27	1,8	3,87	1,46	1,5	1,28
Fe	Zr	1,27	1,8	3,87	1,6	1,4	0,97
Fe	Hf	1,27	1,8	3,87	1,58	1,3	1,01
Fe	Th	1,27	1,8	3,87	1,8	1,1	0,52
Fe	Sc	1,27	1,8	3,87	1,64	1,3	0,68
Fe	Y	1,27	1,8	3,87	1,8	1,2	0,51
Hf	Th	1,58	1,3	1,01	1,8	1,1	0,52
Hf	La	1,58	1,3	1,01	1,88	1,1	0,45
Ir	Y	1,36	2,2	3,6	1,8	1,2	0,51
La	Hf	1,88	1,1	0,45	1,58	1,3	1,01
La	Th	1,88	1,1	0,45	1,8	1,1	0,52
Mo	Zr	1,4	1,8	2,19	1,6	1,4	0,97
Mo	Hf	1,4	1,8	2,19	1,58	1,3	1,01
Mn	Sc	1,3	1,5	3,16	1,64	1,3	0,68
Mn	Y	1,3	1,5	3,16	1,8	1,2	0,51
Mn	Ti	1,3	1,5	3,16	1,46	1,5	1,28
Mn	Zr	1,3	1,5	3,16	1,6	1,4	0,97
Mn	Hf	1,3	1,5	3,16	1,58	1,3	1,01
Mn	Th	1,3	1,5	3,16	1,8	1,1	0,52
Ni	La	1,25	1,8	5,17	1,88	1,1	0,45
Ni	Ti	1,25	1,8	5,17	1,46	1,5	1,28
Ni	Zr	1,25	1,8	5,17	1,6	1,4	0,97
Ni	Hf	1,25	1,8	5,17	1,58	1,3	1,01
Ni	Th	1,25	1,8	5,17	1,8	1,1	0,52
Ni	Sc	1,25	1,8	5,17	1,64	1,3	0,68
Ni	Y	1,25	1,8	5,17	1,8	1,2	0,51
Os	La	1,35	2,2	3,23	1,88	1,1	0,45
Os	Sc	1,35	2,2	3,23	1,64	1,3	0,68
Pd	La	1,38	2,2	3,84	1,88	1,1	0,45
Pt	La	1,39	2,2	3,75	1,88	1,1	0,45
Re	Ti	1,38	1,9	2,69	1,46	1,5	1,28
Re	Zr	1,38	1,9	2,69	1,6	1,4	0,97
Re	Hf	1,38	1,9	2,69	1,58	1,3	1,01
Re	Th	1,38	1,9	2,69	1,8	1,1	0,52
Re	Sc	1,38	1,9	2,69	1,64	1,3	0,68
Ru	Zr	1,34	2,2	3,33	1,6	1,4	0,97
Ru	Hf	1,34	2,2	3,33	1,58	1,3	1,01
Ru	Th	1,34	2,2	3,33	1,8	1,1	0,52
Ru	Y	1,34	2,2	3,33	1,8	1,2	0,51
Ru	La	1,34	2,2	3,33	1,88	1,1	0,45
Ru	Ti	1,34	2,2	3,33	1,46	1,5	1,28
Rh	Sc	1,35	2,2	3,7	1,64	1,3	0,68
Rh	Y	1,35	2,2	3,7	1,8	1,2	0,51
Sc	Y	1,64	1,3	0,68	1,8	1,2	0,51
Sc	La	1,64	1,3	0,68	1,88	1,1	0,45
Sc	Ti	1,64	1,3	0,68	1,46	1,5	1,28

(continued on next page)

– (continued)

A	B	R_A	X_A	Z_A/R_A^3	R_B	X_B	Z_B/R_B^3
Sc	Zr	1,64	1,3	0,68	1,6	1,4	0,97
Sc	Hf	1,64	1,3	0,68	1,58	1,3	1,01
Tc	La	1,36	1,9	2,78	1,88	1,1	0,45
Tc	Ti	1,36	1,9	2,78	1,46	1,5	1,28
Tc	Zr	1,36	1,9	2,78	1,6	1,4	0,97
Tc	Hf	1,36	1,9	2,78	1,58	1,3	1,01
Tc	Th	1,36	1,9	2,78	1,8	1,1	0,52
Ti	Y	1,46	1,5	1,28	1,8	1,2	0,51
Ti	La	1,46	1,5	1,28	1,88	1,1	0,45
Ti	Zr	1,46	1,5	1,28	1,6	1,4	0,97
Ti	Hf	1,46	1,5	1,28	1,58	1,3	1,01
Ti	Th	1,46	1,5	1,28	1,8	1,1	0,52
Y	La	1,8	1,2	0,51	1,88	1,1	0,45
Y	Zr	1,8	1,2	0,51	1,6	1,4	0,97
Y	Hf	1,8	1,2	0,51	1,58	1,3	1,01
Y	Th	1,8	1,2	0,51	1,8	1,1	0,52
W	Zr	1,41	1,7	2,15	1,6	1,4	0,97
Zr	La	1,6	1,4	0,97	1,88	1,1	0,45
Zr	Hf	1,6	1,4	0,97	1,58	1,3	1,01
Zr	Th	1,6	1,4	0,97	1,8	1,1	0,52

REFERENCES

[1] Sudha Priyanga G, Asvini Meenaatci AT, Rajeswara Palanichamy R, Iyakutti K. Comput Mater Sci 2014;84:206–16.
[2] George Lyci, Hrubrak Ross, Rajan Krishna, Saxena Surendra K. J Alloys Compd 2009;478:731–5.
[3] Bououdina M, Grant D, Walker G. Int J Hydrogen Energy 2006;31:177–82.
[4] Sakintuna B, Lamari-Dakrim F, Hirscher M. Int J Hydrogen Energy 2007;32:1121–82.
[5] Gasiorowski A, Iwasieczko W, Skoryna D, Drulis H, Jurczyk M. J.Alloys Compd 2004;364(1–2):283–8.
[6] Benyelloul K, Bouhadda Y, Bououdina M, Faraoun HI, Aourag H, Seddik L. Int J Hydrogen Energy 2014;39:12667–75.
[7] Rusman NAA, Dahari M. Int J Hydrogen Energy 2016;41:12108–26.
[8] Okada M, Kuriiwa T, Kamegawa A, Takamura H. Sci Eng A 2002;329–331:305–12.
[9] Wang J, Shang S-L, Wang Y, Mei Z-G, Liang Y-F, Du Y, et al. Comput Coupling Phase Diagr Thermochem 2011;35:562–73.
[10] Herbst JF. J Alloys Compd 2002;337:99–107.
[11] Griessen R, Driessen A, De Groot DG. J Less Common Metals 1984;103:235–44.
[12] Matysik P, Czujko T, Varin RA. Int J Hydrogen Energy 2004;39:398–405.
[13] de Boer FR, Boom R, Mattens WCM, Miedema AR, Niessen AK. In: de Boer FR, Pettifor DG, editors. Cohesion and structure, vol. 1. Amsterdam: North-Holland; 1988.
[14] Klaveness A, Fjellvag H, Kjekshus A, Ravindran P, Swang O. J Alloys Compd 2009;469:617–22.
[15] Watson RC, Bennett LH. Calphad 1981;5:25–40.
[16] Van Mal HH. Philips Res Repts 1976;(Suppl. 1):61.
[17] Kagami N, Wamoto RI, Tani T. Fuel 2005;84:279–85.
[18] Benyelloul K, Seddik L, Bouhadda Y, Bououdina M, Fenineche N, Aourag H, et al. Int J Hydrogen Energy 2016;41:11254–63.
[19] Benyelloul K, Aourag H. Comp Mat Sci 2013;67:353–8.
[20] Benyelloul K, Aourag H. Comp Mat Sci 2013;77:330–4.
[21] Guo Jin, Wei Wen-lou, Ma Shu-yuan. J Alloys Compd 2004;372:136–40.
[22] Shih Jing-Shiang, Tzeng Yih-Fong, Yang Jin-Bin. Mater Des 2011;32:1253–61.
[23] Gu Fu, Hall Philip, Nicholas JMiles, Ding Qianwen, Wu Tao. Mater Des 2014;62:189–98.
[24] Wang Cheng, Wen Miao, Bai Lihong, Zhang Tong. Optik 2016;127:1007–10.
[25] Fung HC, Kang PC. J Mater Process Technol 2005;170:602–10.
[26] Tumac D. Measurement 2016;80:12–20.
[27] Morgan D, Ceder G, Yip S, editors. Handbook of materials modelling. Netherlands: Springer; 2005. p. 395–421.
[28] Haghdadi N, Zarei-Hanzaki A, Khalesian AR, Abedi HR. Mater Des 2013;49:386–91.
[29] Dias JLR. N Silvertre Eng Struct 2011;33:2015–24.
[30] Kittel C. Introduction to solid state physics. 7th ed. John Wiley & Son. Inc; 1996.
[31] Matar SF. Prog Solid State Chem 2010;38:1–37.
[32] Acharya N, Fatima B, Chouhan SS, Sanyal SP. Compt Mat Sci 2015;98:226–33.
[33] Hirscher Michael, editor. Handbook of hydrogen storage: new materials for future energy storage. Weinheim: Wiley-VCH Verlag GmbH & KGaH; 2010.
[34] Soubeyroux JL, Fruchart D, Lorthioir G, Ochin P, Colin D. J Alloys Compd. 1993;196:127–32.
[35] Ciric KD, Kocjan A, Gradisek A, Koteski VJ, Kalijadis AM, Ivanovski VN, et al. Int J Hydrogen Energy 2012;37(10):8408–17.
[36] Newkrik HW. A literature survey of metallic ternary and quaternary hydrides, vol. 51244. UCRI; 1974.
[37] Korst WL. Rpt NAA-SR-6881. May 1962.
[38] Buschow KHJ. In: Gschneidner Jr KA, Eyring L, editors. Handbook on the physics and chemistry of rare earths, vol. 6. Amsterdam: North-Holland; 1984. p. 1.
[39] Griessen R, Riesterer T. In: Schlapbach L, editor. Topics in applied physics, vol. 63. Berlin: Springer; 1988. p. 219.