Statistically based assessment of formation enthalpy for intermetallic compounds

R.F. Zhang$^{a,b,*}$, K. Rajan$^b$

$^a$ School of Materials Science and Engineering and International Research Institute for Multidisciplinary Science, Beihang University, Beijing 100191, PR China

$^b$ Institute for Combinatorial Discovery and Department of Materials Science and Engineering, Iowa State University, Ames, IA 50011, USA

**Abstract**

A simplified composition-independent atomic size difference factor is proposed to account for the statistical deviation of original Miedema's model for formation enthalpies of binary intermetallic compounds from experimental data. We demonstrate that with the incorporation of this new factor, one can achieve a high level of correlation with respect to available experimental data. Using this statistically refined model, we propose a methodology for assessing where the next level of statistical refinement is needed, utilizing a data-mapping framework of ‘critical descriptors’. This provides a way to identify the search space where needs further experimental work to be performed.

© 2014 Elsevier B.V. All rights reserved.

A decade ago, by comparing the experimental values published by Kleppa et al. and Meschel et al. [1–6] with the calculated ones from original Miedema’s theory, it was found interestingly that over 90% of the calculated values are lower or more negative than the experimentally determined ones [7,8]. Such systematical deviation arose much interest and attention in materials community, and accordingly some modification schemes (see Ref. [8] and related references cited therein) have been proposed to improve its predictability. Among these studies, the atomic size difference (ASD) factor proposed by Zhang et al. [7] has been generally believed to play an important role in affecting the precision of the calculation, yet was not considered in the original Miedema’s theory.

The introduction of ASD factor is not trivial, but based on the fact that the predicted values with large deviation from experiments always appears in the binary transition-metal systems featuring with large atomic size difference [7]. In order to account for it, ASD factor is proposed to be the intrinsic origin and included into Miedema’s model, and consequently its predictability is much improved indeed, demonstrating the success of discovering the underlying physics from intuitive data mining.

In the following years, several groups tried different version of modification of ASD factor to further improve the model predictability (e.g. [8]), unfortunately no significant progress has been obtained so far due to the restriction of the form modification in modeling, not to a discovery of new physics. In this letter, we develop a parameterization scheme for a compositional independent atomic size difference factor in Miedema’s model. We use this new improved model to map a continuous phase space of enthalpy data, by using data interpolation methods. By then comparing our new statistical refinement with experimental data, we demonstrate that we have a robust model over a wide composition space and at the same time identify specific regions where refinements can still be made.

According to Miedema’s theory, the standard formation enthalpy of the intermetallic compounds consisting of transition, lanthanides and actinide metals, can be evaluated by two critical parameters, i.e. the electronegativity difference and the electron-density discontinuity. The electronegativity difference is expressed by modified work function \( \Delta \Phi^\ast \) between two constituent metals is responsible for a negative contribution to the formation energy/enthalpy. It provides a major driving force for the formation of an intermetallic compound by the charge transfer, which decreases the contact potential difference between two dissimilar metals. This parameter can be expressed by the quadratic difference of work function of two constituent metals. The discontinuity of the electron density at the boundaries of the dissimilar Wigner–Seitz (WS) cells is defined as \( \Delta n^{1/2} \), and the formation enthalpy of an intermetallic compound can be expressed by [9–12],

\[
\Delta H_{\text{form}}^{\text{tot}} = c_1 \cdot c_2 \cdot (1 + 8(c_3' \cdot c_4')^2) \frac{c_3 \cdot V^{1/3} + c_4 \cdot V_{\ast}^{2/3}}{A_{\ast}^{-1/3}} \left( -P(\Delta \Phi^\ast)^2 + q(A^{1/2} \cdot \Delta n^{1/2})^2 \right)
\]

**References**


http://dx.doi.org/10.1016/j.cplett.2014.08.024 0009-2614/© 2014 Elsevier B.V. All rights reserved.
where, \( c_A^i = c_{AV}^2/\sqrt{1 + c_{AV}^2/3 + c_{B}^2/3} \), \( c_B^i = c_{BV}^2/\sqrt{1 + c_{BV}^2/3 + c_{AV}^2/3} \), \( n_i^{ws} = (n_i^{wsA} + n_i^{wsB})/2 \), \( c_i \) and \( v_i \) (\( i = A \) or \( B \)) are the atomic concentration and molar volume of component \( i \) respectively (where the molar volumes account for volume changes during alloying due to charge transfer [11,12]). \( \Delta \Phi^* \) is the electronegativity difference parameter or chemical potential difference, and \( \Delta n_{ws}^c \) is the electronic density discontinuity at the boundary of the Wigner–Seitz cell. \( P, Q \) are the empirical parameters and their ratio \(QP=9.4 \) is a critical value which makes \( (\Delta \Phi^*)^2 \) versus \( (\Delta n_{ws}^c)^2 \) plot separate the positive and negative alloy systems [9–12].

In order to account for the statistical deviation of the calculated values by original Miedema’s model from the experimental ones, a third parameter, i.e. the atomic size difference (ASD) factor [7], was introduced into the modeling because it would frequently lower the contact area between the two dissimilar WS cells and decrease the binding energy between the two dissimilar atoms. In addition, the atomic size difference factor was previously proposed by Hume-Rothery et al. [13] as one of the critical parameters in determining the formation of intermetallic compounds, which however, was not explicitly considered in the original Miedema’s model. Accordingly, the modified version can be expressed in a general form,

\[
\Delta H_m = Sc \cdot \Delta H_{Mied} \tag{2}
\]

where, \( Sc \) is the ASD factor which was originally introduced by Zhang et al. [7] to account for the systematic deviation of calculated values from experiments. It was originally proposed to be dependent on two fundamental parameters such as the atomic sizes of constituent metals and the composition of solute components, indicating its composition-dependent feature. Notice that one must distinguish the solute and solvent atoms when applying the previously proposed ASD by Zhang et al. [7]. Since the composition-dependence cannot be considered as an intrinsic origin of the systematic deviation, we may remove it and consequently the reformulated model is expressed as,

\[
\Delta H_m = \frac{\alpha \cdot V_A^{2/3} / V_B^{2/3}}{(V_A^{2/3} + V_B^{2/3})^2} \cdot \Delta H_{Mied} \tag{3}
\]

or in its full form as

\[
\Delta H_m = \frac{\alpha \cdot V_A^{2/3} / V_B^{2/3}}{(V_A^{2/3} + V_B^{2/3})^2} \cdot \frac{c_A^i \cdot c_B^i \cdot (1 + 8(c_A^i \cdot c_B^i)^2)}{c_A^i \cdot c_B^i} \cdot \frac{c_A^i V_A^{2/3} + c_B^i V_B^{2/3}}{(n_i^{ws}^{1/3})_{av}}
\]

\[
\times [-P(\Delta \Phi^*)^2 + Q(\Delta n_{ws}^c)^2] \tag{4}
\]

where, \( \alpha \) is an empirical parameter which is determined by a statistical data-driven regression. Because of one single undetermined parameter, linear regression is used to model the relationship between a dependent variable and one or more explanatory variables. The case of one explanatory variable is called simple linear regression. Consequently, the empirical parameter \( \alpha \) is determined as \( \alpha = 3.2598 \) (\( P=14.2 \)) or \( \alpha = 3.2829 \) (\( P=14.1 \)) for the intermetallic compounds consisting of transition, lanthanides and actinides metals. \( Sc = \alpha \cdot V_A^{2/3} / V_B^{2/3} \) can be expressed in a more physically understandable way as \( Sc = \alpha \cdot (1 - (V_A^{2/3} - V_B^{2/3})^2) / (V_A^{2/3} + V_B^{2/3})^2 = \alpha \cdot V_A^{2/3} / (V_A^{2/3} + V_B^{2/3})^2 \)

It can be seen that the ASD factor, i.e. \( Sc \), reflects in fact the magnitude of contact surface mismatch during alloying, and is an intrinsic properties of a binary alloy system when dissimilar atoms combines together since it does not depend on the composition of solute anymore. More expressive is its simplified form in a similar manner to electronic density misfit \( \Delta n_{ws}^c \) and electronegativity difference \( \Delta \Phi^* \). In addition, we have tried different nonlinear form of ASD besides of the linear regression. For example, assuming that \( \chi = \frac{V_B^{2/3} / V_A^{2/3}}{V_B^{2/3} + V_A^{2/3}} \) and \( Sc = \alpha \chi + \beta \chi^2 \), we can get a similar good statistical agreement with experiments (as will be shown below) too, however, a negative regressed value of \( \beta \) without any constraints will result in an unphysical trend of ASD.
effect. In such case it can be noticed that the ASD factor may be adopted in different form but must abide by a general increasing trend with a reasonable physical picture.

To indicate the improvement of the present model, Figure 1 shows the relationship between experimentally determined data and the calculated standard formation enthalpies of 435 intermetallic compounds which are collected from a series of experimental reports [1–6], by original Miedema’s model [9–12], Zhang’s previously modified version [7] and the presently proposed statistical model. In these figures, the X-axis is the experimental values and Y coordinate stands for the calculated values by the three models. Two dashed lines are set to define a data zone with an error bar of ±30 kJ/mol. It can be seen from Figure 1 that over 99% circle symbols (by Zhang’s previously modified version) or star symbols (by presently proposed statistical model) are within the data zone whereas about 95% square symbols (by original Miedema’s model) are located within the same zone. To qualitatively show the improvement, we have calculated the standard deviation between experimental and predicted values, which is defined as $\sigma = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} (y_i - \bar{y})^2}$, where $y_i$ and $\bar{y}$ are the predicted and experimental values respectively. The calculated $\sigma$ are 19.08 for original Miedema’s model, 13.38 for Zhang’s previously modified version and 12.51 for the present statistical model. This comparison suggests that both Zhang’s previously modified model and present statistical model can significantly improve the model predictability, but more significant is that the removal of composition-dependence from the model makes the scatter data points much condensed around the diagonal straight lines. Noted that some improvement may also be obtained by the modification of model parameters of electronic density and electronegativity as previously used by Chen et al. [14] and Zhang et al. [7]. However, such parameter optimization might be arbitrary and intuitive [14], and a further physical validation is much necessary according to Miedema’s scheme of statistical classification of positive and negative systems [11,12]. In our present model, all the model parameters are adopted from the original textbook by Miedema and reproduced by Bakker without any modification [11,12].

To further illustrate improvement of the present simplified model, Table 1 lists the calculated formation enthalpies of some compounds by the present model and by the original Miedema’s model to compare with the experimental values for those which features a large atomic size difference. From the table, one may see that the present model can considerably improve the calculation precision, comparing with original Miedema’s model. For more detailed comparison, we may refer to our freely released software: Miedema Calculator Ver.3.5 [15], in which any binary combination may be easily calculated based on original Miedema’s model, Zhang’s previous model and present statistical model.
Table 1
Comparison of the calculated formation enthalpies by different models and the experimental ones for some representative intermetallic compounds with 1:1 stoichiometry featuring large atomic size difference. Unit is kJ/mol of atoms. Experiments: experimental values by Kleppa et al. [1–6]; present: present statistical model; previous: Zhang’s previous model [7]; Miedema: original Miedema’s model [11–12].

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Present</th>
<th>Experiments</th>
<th>Previous</th>
<th>Miedema</th>
</tr>
</thead>
<tbody>
<tr>
<td>AgLa</td>
<td>−33.8</td>
<td>−16.9</td>
<td>−32.1</td>
<td>−43.2</td>
</tr>
<tr>
<td>AgY</td>
<td>−34.6</td>
<td>−26.8</td>
<td>−34.1</td>
<td>−43.6</td>
</tr>
<tr>
<td>AgSc</td>
<td>−34.5</td>
<td>−26.2</td>
<td>−37.1</td>
<td>−42.5</td>
</tr>
<tr>
<td>AgZr</td>
<td>−24.9</td>
<td>−3.1</td>
<td>−27.5</td>
<td>−30.7</td>
</tr>
<tr>
<td>AgHf</td>
<td>−15.4</td>
<td>−10.7</td>
<td>−17.2</td>
<td>−18.9</td>
</tr>
<tr>
<td>AuLa</td>
<td>−85.6</td>
<td>−69.9</td>
<td>−80.1</td>
<td>−108.0</td>
</tr>
<tr>
<td>AuY</td>
<td>−88.4</td>
<td>−78.7</td>
<td>−86.1</td>
<td>−110.2</td>
</tr>
<tr>
<td>AuSc</td>
<td>−90.7</td>
<td>−76.1</td>
<td>−97.1</td>
<td>−111.4</td>
</tr>
<tr>
<td>AuHf</td>
<td>−78.0</td>
<td>−57.0</td>
<td>−87.0</td>
<td>−95.8</td>
</tr>
<tr>
<td>PtCe</td>
<td>−93.7</td>
<td>−103.8</td>
<td>−86.1</td>
<td>−119.6</td>
</tr>
<tr>
<td>PtSc</td>
<td>−107.1</td>
<td>−104.8</td>
<td>−110.2</td>
<td>−132.1</td>
</tr>
<tr>
<td>PtHf</td>
<td>−109.0</td>
<td>−113.7</td>
<td>−117.0</td>
<td>−134.4</td>
</tr>
<tr>
<td>PtY</td>
<td>−97.3</td>
<td>−104.0</td>
<td>−91.5</td>
<td>−122.9</td>
</tr>
<tr>
<td>YRh</td>
<td>−61.7</td>
<td>−76.1</td>
<td>−56.8</td>
<td>−79.2</td>
</tr>
<tr>
<td>YPd</td>
<td>−97.0</td>
<td>−94.9</td>
<td>−90.9</td>
<td>−123.2</td>
</tr>
<tr>
<td>YIr</td>
<td>−61.1</td>
<td>−65.9</td>
<td>−56.4</td>
<td>−77.9</td>
</tr>
<tr>
<td>PdSc</td>
<td>−103.7</td>
<td>−85.3</td>
<td>−106.1</td>
<td>−128.3</td>
</tr>
<tr>
<td>PdZr</td>
<td>−110.1</td>
<td>−66.1</td>
<td>−115.8</td>
<td>−136.3</td>
</tr>
<tr>
<td>RhLa</td>
<td>−55.2</td>
<td>−56.3</td>
<td>−49.0</td>
<td>−72.2</td>
</tr>
<tr>
<td>HfRe</td>
<td>−36.3</td>
<td>−37.0</td>
<td>−38.7</td>
<td>−44.9</td>
</tr>
</tbody>
</table>

Figure 3. Ternary property difference landscape maps to indicate the difference between the calculated formation enthalpies and experimental ones, (a) original Miedema’s model [9–12], (b) Zhang’s previously modified version [7] and (c) the presently proposed statistical model, with the 0–1 scaling of electronegativity difference (END), electron density discontinuity (EDD), and atomic size difference (ASD). (For interpretation of the references to color in figure legend, the reader is referred to the web version of the article.)

To rapidly test the overall robustness of our refinements, we use a data descriptor approach to [16,17] study the Miedema’s model where we explore how well our predictions correspond to experiments in the context of key metrics (electronegativity difference, electron density discontinuity, and atomic size difference) that influence the Miedema model rather than explore every composition. The impact of predictive models in creating a virtual library of compounds and identifying regions of interest or high impact can be appreciated considering the ternary property landscape maps which depict the response of an independent variable to changes in a mixture of three components. Figure 2 shows the ternary property landscape map to indicate the distribution of experimental formation enthalpies, the calculated formation enthalpies by original Miedema’s model and by present statistical model with the scaling of electronegativity difference (END), electron density difference (EDD), and atomic size difference (ASD). END and EDD stand for \(-P(\Delta \Phi)^2/(\Delta n_{ws}^{-1/3})_{av}\) and \(Q(\Delta n_{ws}^{-1/3})_{av}/\Delta n_{ws}^{-1/3}\) respectively, while ASD indicates the other parts as shown in Eq. (4). A common intersection method can be used to determine the ratios of the three components, i.e. a straight line is drawn from each corner, through the point of interest, to the opposite side of the triangle. The ratios can be determined by dividing the lengths of segments by the entire corresponding line.
The plot maps the simultaneous effect of END, EDD and ASD on the formation enthalpies, as an extrapolation about the existing sparse data. Here the values of 435 formation enthalpies are entered into ternary property landscape map and a triangulation connection method is employed to determine contours around the existing data, enclosing regions of the same formation enthalpy. A linear interpolation is used for the intersection points of contour lines and triangle sides. After tracing all the characteristic points, the smoothing of contours is performed by the B-Spline curve fitting scheme. All details can be found in Refs. [18,19]. The axes are scaled from 0 to 1 in a range consistent with the minimum and maximum values of parameter space respectively. The regions beyond the boundaries of the landscape reflect either the unavailability of compounds with these combinations of axes parameters or may identify realistic design limits in formation enthalpies. Similarly within the landscape itself there exist gaps in the data reflecting potential discovery of new compounds. The ternary property landscape map provide also a unique way to identify the design limits for intermetallic compounds and the unexplored gaps in the existing data space, serving as a guide to regions of unexpected high thermal stability with negative formation enthalpy.

The statistical improvement can also be compared by identifying the difference between the calculated formation enthalpies and experimental ones. Figure 3 shows the ternary property difference landscape maps to indicate the difference between the calculated formation enthalpies and experimental ones, i.e. the original Miedema’s model [9–12], Zhang’s previously modified version [7] and the present statistical model. As can be seen, large deviation from experimental data is found for original Miedema’s model, much improved by the Zhang’s previously modified model. More impressively, the present statistical model can further reduce the disagreement magnitude (see the red regimes shown in the figures), indicating a statistical method can efficiently improve the model predictability.

In summary, we have proposed a new and improved atomic size difference (ASD) factor as an intrinsic model parameter for predicting formation enthalpies of binary transition metal intermetallic compounds by means of statistical learning scheme. The newly proposed model rules out the composition-dependence of ASD factor. Consequently three intrinsic components are included in the improved statistical model: the electronegativity difference $\Delta \Phi^*$, the electron density discontinuity $\Delta n_{\text{dis}}^{1/3}$ and the ASD factor $Sc$ between two constituent metals. The physical validation of the three parameters is explored in terms of a ternary property landscape mapping technique which confirmed the newly proposed model not only statistically reproduce the experimental data very well, but provide a way to identify where even further refinements via experiments may be targeted.

**Acknowledgements**

The authors gratefully acknowledge support from the Air Force Office of Scientific Research Grant # FA9550-12-0496 and National Science Foundation Grant # DMS-13-07811. R.FZ would also like to thank the Zhuo-Yue Hundred Talents Plan of Beihang University and National Thousand Young Talents Program of China. KR acknowledges support from the Wilkinson Professorship of Interdisciplinary Engineering.

**References**

[15] Miedema Calculator Ver. 3.5 software with functional of present model can be freely accessed via link: https://sites.google.com/site/miedemacalculator or send e-mail to authors.