

# Linking Length Scales via Materials Informatics

Zi-Kui Liu, Long-Qing Chen, and Krishna Rajan

*Seeking structure-property relationships is an accepted paradigm in materials science, yet these relationships are often not linear. The challenge is to be able to link materials behavior and seek patterns among multiple length and time scales. Materials informatics and statistical learning techniques permit one to survey complex, multiscale information in an accelerated and statistically robust and yet meaningful manner. When this is coupled with advanced tools for computational thermodynamics and kinetic simulations, the result is a powerful computational infrastructure for materials design. This paper provides an overview of the value of this integrated approach to materials modeling in addressing the challenges in linking length scales.*

## INTRODUCTION

Materials design is a process to determine optimal combinations of material chemistry, processing routes, and processing parameters to robustly meet specific performance requirements such as mechanical properties and corrosion resistance. This process is iterative by nature due to the incompleteness of the design knowledge base and the lack of one-to-one correspondence in this inverse problem. This paper provides two computational approaches to materials design. One is to apply a set of computational strategies that capture different length scales of materials behavior and integrate information across these length scales using advanced discretization

and parallel algorithms and a software architecture for distributed computing systems. This is termed "hard modeling," and builds on well-established constitutive relationships.

The other approach is to seek, through statistical learning techniques, heuristic relationships between data across length scales. Of course, the interpretation of such relationships has to be built on fundamental materials-science principles. This approach can significantly accelerate the process of screening computational results. While a fundamental tenet in materials science is to establish structure-property relationships, it is the life sciences and organic chemistry community that has formally introduced the concept of quantitative

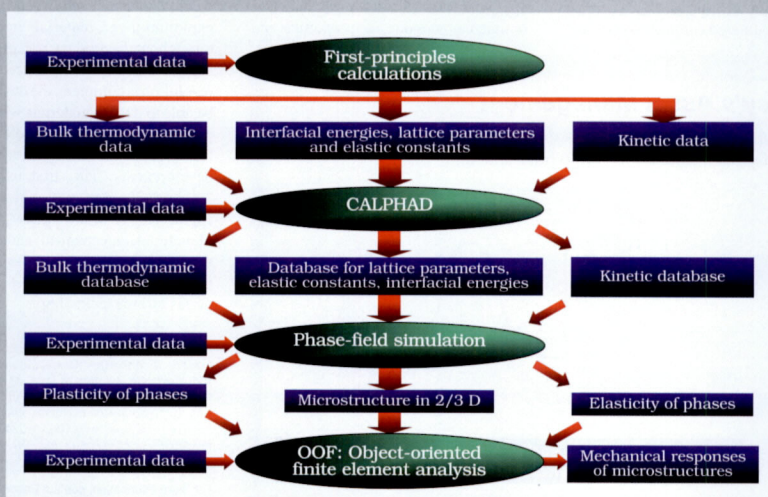


Figure 1. An integrated four-stage multiscale approach for multicomponent materials modeling and simulation.

structure activity (also termed property) relationships (QSAR or QSPR). Unlike classical materials-science approaches, which relate structure and function through physically based models (e.g., thermochemical and microstructural evolution models), QSARs are derived from a model-independent approach, sometimes called "soft modeling."

## HARD MODELING STRATEGY

Recently, an integrated framework was developed for multi-scale materials simulation and design<sup>1</sup> named MatCASE (Materials Computation and Simulation Environments), which involves four major computational steps as shown in Figure 1: atomic-scale first-principles calculations to predict thermodynamic properties, lattice parameters, and kinetic data of unary, binary, and ternary compounds and solutions phases; the CALPHAD (CALCulation of PHase Diagrams) data-optimization approach to model thermodynamic properties, lattice parameters, and kinetic data of multicomponent systems; a multicomponent phase-field approach to predict the evolution of microstructures in one to three dimensions; and finite-element analysis to generate the mechanical response from simulated microstructures. These four stages are integrated with advanced discretization and parallel algorithms and a software architecture for distributed computing systems. This paper will focus on microstructure evolutions of nickel-based superalloys starting from the generation of materials data needed for such simulations and their integration in terms of information flow and data processing.

### First-Principles Calculations of Materials Properties

First-principles calculations, based on density functional theory, require only knowledge of the atomic species and crystal structure. They yield quantities related to the electronic structure and total energy of a given system and can be used to accurately predict phase stabilities of compounds at 0 K. Coupling with frozen phonon or linear response techniques, it opens the possibility for exploring finite-temperature vibrational effects for both stable and metastable phases. The latter is often crucial to

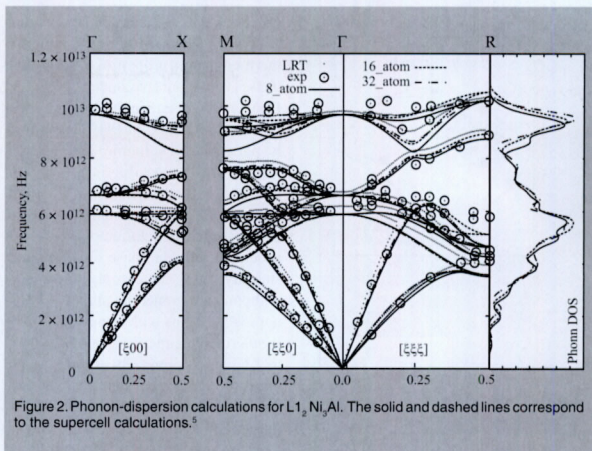


Figure 2. Phonon-dispersion calculations for  $L1_2$   $Ni_3Al$ . The solid and dashed lines correspond to the supercell calculations.<sup>5</sup>

materials properties but difficult to isolate and study experimentally.

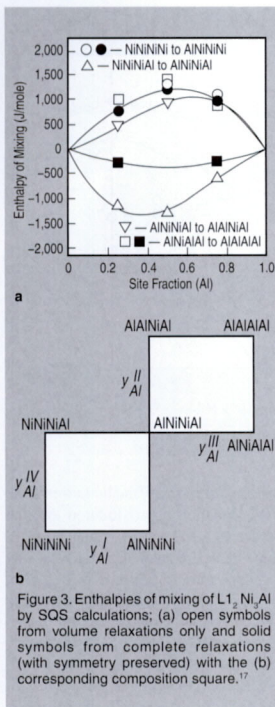
For thermodynamic properties, there are three typically additive contributions to the free energy. The first contribution is the cold energy or the 0 K total energy. In this case, the atoms are kept fixed at their static lattice positions. Second, for thermodynamic properties at finite temperatures, the contribution of lattice thermal vibration needs to be taken into account. Theoretically, the commonly accepted method is the lattice dynamics or phonon approach. When temperature is increased, especially for cases when the electronic density of state at the Fermi level is high, the third contribution to be included is the thermal electronic contribution (TEC). When the magnetic contribution and the electron-phonon interactions are neglected, TEC to the free energy  $F_{el}$  is obtained from the energy and entropy contributions (i.e.,  $E_{el} - TS_{el}$ ). This approach enables the study of the phase stability at high temperatures of pure elements and compounds.<sup>2,3</sup>

The current first-principles implementations of phonon theory are divided into two categories: the linear-response method and the supercell method. In the linear-response method, the normal frequencies (i.e., phonon frequencies) associated with microscopic displacements of atoms in a crystal are calculated by means of dynamical matrix which is the reciprocal-space expression (Fourier transform) of interatomic force constant

matrix. The supercell method adopted the frozen-phonon approximation through which the changes in total energy or forces are calculated in the real space by displacing the atoms from their equilibrium positions. If one is only interested in the gross phonon frequencies, the results of supercell calculation and linear-response calculation are very similar.<sup>4,5</sup> The Alloy-Theoretic Automated Toolkit (ATAT) code developed by van de Walle et al.<sup>6</sup> serves as the interface connecting to the Vienna ab-initio simulation package (VASP) code for phonon calculations in the supercell method.

While the calculation of the electronic structure of perfectly ordered periodic structures with first-principles methods is well established, it is not the case for phases with finite solubilities. A couple of representative methods used to obtain the properties of random solid solutions from theoretical calculations include the coherent potential approximation,<sup>7</sup> the cluster expansion,<sup>8,9</sup> and the special quasi-random structures (SQS).<sup>10</sup> Each method has its own limitations in representing random solutions, and it seems that the SQS provide an optimal combination in terms of computational efficiency and accuracy.<sup>11</sup> The SQS mimic a random solution phase by creating a small (4–48 atoms) periodic structure that best satisfies the pair and multisite correlation functions corresponding to a random solid solution, up to a certain coordination shell. In terms of computa-





tional efficiency, one single calculation on SQS can give various properties of random alloys, and only the SQS allow direct calculations of real-space quantities (e.g., lattice parameters and bond lengths) of alloys in comparison with the other approaches. Using the ATAT code,<sup>6</sup> the SQS are now available for body-centered cubic,<sup>12</sup> B2,<sup>13</sup> Laves phases,<sup>14</sup> halite,<sup>15</sup> hexagonal-close packed,<sup>16</sup> and L<sub>12</sub> structures.<sup>17</sup>

Additionally, first-principles calculations of a wide range of materials properties such as interfacial energy,<sup>18,19</sup> antiphase boundary energy,<sup>20</sup> diffusivity,<sup>21,22</sup> and elastic constants,<sup>23</sup> are also possible.

### CALPHAD Modeling of Materials Properties

The CALPHAD approach was pioneered by L. Kaufman<sup>24</sup> for modeling of thermodynamic properties and has been extended to modeling the atomic mobility, molar volume, and other materials

properties. It is based on mathematically formulated models describing the thermodynamic properties of individual phases with model parameters evaluated from thermochemical and phase equilibrium data. The CALPHAD approach is particularly valuable in materials science and engineering in comparison with physics and chemistry due to more complicated systems studied such as multicomponent solution phases. It has produced reliable phase diagrams and stability maps for complicated multicomponent commercial alloys.<sup>25,26</sup> This approach begins with the evaluation of the thermodynamic descriptions of unary and binary systems. By combining the constitutive binary systems with ternary data, ternary interactions and Gibbs energy of ternary phases are obtained.

Models for the Gibbs energy are based on the crystal structures of the phases. For pure elements and stoichiometric compounds, the most commonly used model is the one suggested by the Scientific Group Thermodata Europe (SGTE).<sup>27</sup> The SGTE data for pure elements have been compiled by A. Dinsdale.<sup>28</sup> For multicomponent solution phases, the molar Gibbs energy has the general formula shown in Equation 1 (all equations are shown in the Equations table), where  ${}^{\text{ref}}G_m$  represents the reference Gibbs energy of the phase,  ${}^{\text{ideal}}G_m$  the ideal mixing contribution, and  ${}^{\text{ex}}G_m$  the excess Gibbs energy of mixing due to non-ideal interactions. A variety of mathematical models have been developed for these three terms.<sup>29</sup> For phases experiencing order-disorder transitions, one may separate the ordering contribution from the Gibbs energy of a disordered phase so the disordered phase can be modeled independently, as suggested by I. Ansara et al.<sup>30</sup>

The atomic mobility ( $M_i$ ) is related to the tracer diffusion coefficient ( $D_i^*$ ) through the Einstein Equation 2, with  $D_i^*$  expressed by  $D_i^* = D_i^0 \exp(-Q_i/RT)$ , where  $D_i^0$  and  $Q_i$  are the prefactor and activation energy. The atomic mobility can be written as Equation 3, with  $\Delta\Phi_i = Q_i - RT \ln D_i^0$ . The atomic mobility can be modeled using the CALPHAD approach, similar to thermodynamics.<sup>31</sup> The atomic mobility can be used to calculate the intrinsic diffusivity in the lattice-fixed frame of reference and the interdiffusion coefficients in the volume-

fixed frame of Reference 31. The effect of chemical ordering can be treated using a two-sublattice model by T. Helander and J. Ågren.<sup>32</sup>

The molar volume of a phase can be calculated from its molar Gibbs energy provided the pressure dependence of Gibbs energy is available.<sup>33</sup> As the effect of pressure on Gibbs energy under ambient conditions is relatively small with respect to temperature and compositions, alternatively, one can model the lattice parameter directly using the similar phenomenological approach as recently carried out by T. Wang et al.<sup>34</sup>

### Phase-Field Simulation Principles

In a phase-field model, a microstructure is described using one or more physical and/or artificial field variables.<sup>35</sup> These variables are uniform inside a phase or domain away from the interfaces. The same phase or the same types of domains have the same values for the field variables. Different values of the field variables, for example, 0 and 1, distinguish different phases or domains. Across the interfaces between different phases or domains, the field variables vary continuously from one uniform value corresponding to one type of phase or domain to another corresponding to another phase or domain. Therefore, the interfaces in a phase-field model are diffuse and possess a certain thickness. Field variables can be either conserved or non-conserved, depending on if they satisfy the local conservation law,  $\partial\phi/\partial t = \nabla \cdot \mathbf{J}$  where  $\phi$  is a field variable and  $\mathbf{J}$  is the corresponding flux. For example, composition and temperature fields are both conserved while long-range order parameter fields describing ordered domain structures are non-conserved.

In the phase-field approach, the thermodynamics of an inhomogeneous microstructure are described by the diffuse-interface theory of J. Cahn and J. Hilliard.<sup>36</sup> The temporal and spatial evolution of the field variables follows the same set of kinetic equations. All conserved fields,  $c_i$ , evolve with time according to the Cahn-Hilliard equation,<sup>37</sup> or simply the diffusion equation in the case that no gradient energy is introduced for the conserved variable. The non-conserved fields,  $\eta_i$ , are governed by the Allen-Cahn equation.<sup>38</sup> The

microstructure evolution of a system can be obtained by numerically solving the systems of Cahn-Hilliard diffusion equations and the Allen-Cahn relaxation equations subject to appropriate initial and boundary conditions.

In order to relate the phase-field parameters to the experimentally measurable thermodynamic and kinetic properties, one has to examine the phase-field equations in the sharp- and/or thin-interface limit. This is particularly true for phase-field models with artificial field variables for which the corresponding kinetic parameters are not directly related to the measurable physical properties. A sharp-interface analysis<sup>39</sup> matches the phase-field parameters at the limit of zero interfacial thickness to experimentally measured thermodynamic and kinetic properties while a thin-interface analysis<sup>40–42</sup> allows the variation of the phase-field variable over a certain thickness for the interface. It is shown by A. Karma<sup>40</sup> that a phase-field simulation using the thin-interface asymptotics permits one to use a larger interface width and thus larger grid size.

#### Applications to the Ni-Al and Ni-Mo Systems

Nickel-based superalloys consist of ordered intermetallic  $\gamma'$ -L1<sub>2</sub> precipitates embedded in a disordered face-centered cubic (fcc)  $\gamma$ -fcc matrix. The  $\gamma'$  precipitate morphology and spatial distribution are known to depend on a number of factors including the alloy compositions, temperature, presence of dislocations, and applied stress direction. Therefore, extensive efforts have been devoted to understanding the evolution of the  $\gamma'$  precipitate microstructure, both experimentally and theoretically, over the past 50 years.<sup>43</sup> This section presents recent work on Ni-Al through combining first-principles calculations, CALPHAD modeling, and phase-field simulations to determine the relationships between chemistry and microstructure at high temperatures.<sup>44</sup> The ternary Ni-Al-Mo system is one of the basic systems for nickel-based superalloys.<sup>45</sup>

First-principles calculations were performed using VASP<sup>46</sup> with the ultrasoft pseudopotentials and the generalized gradient approximation.<sup>47</sup> Spin-polarized configurations are used for checking the magnetism of nickel and nickel-rich

compounds. The thermodynamic properties of aluminum, nickel, NiAl, and Ni<sub>3</sub>Al can be systematically calculated by means of the linear-response theory (LRT)<sup>4</sup> and the supercell method (SC)<sup>5</sup> with the phonon dispersion calculations for L1<sub>2</sub> Ni<sub>3</sub>Al shown in Figure 2. For the L1<sub>2</sub> structure represented by a four-sublattice model (i.e., (Al,Ni)(Al,Ni)(Al,Ni)(Al,Ni)), the SQS for the random mixing in each sublattice are equivalent (i.e., (Al,Ni)(i)(j)(k) where i, j, and k are either aluminum or nickel). A 64-atom SQS<sup>17</sup> was developed using ATAT. The calculated enthalpies of formation using the ultrasoft pseudopotentials for four combinations of i, j, and k are shown in Figure 3. There is an equal number of combinations where aluminum substitutions take place in different sublattices. It is observed that the atomic environments have significant impacts on the enthalpies of mixing.

For lattice distortion and lattice parameters, 108-atom supercells are employed with one solute atom in each supercell to study the lattice distortions caused by solute atoms.<sup>44</sup> Eleven commonly used alloying elements in nickel-based alloys were chosen, namely, Al, Co, Cr, Hf, Mo, Nb, Re, Ru, Ta, Ti, and W. The lattice distortions due to solute atoms can be separated into two categories, local lattice distortion and macroscopic lattice distortion, with the latter represented

by the overall lattice parameter change represented by  $\Delta a = a_{\text{sol}} - a_{\text{pure}}$  with  $a_{\text{pure}}$  and  $a_{\text{sol}}$  being the lattice parameters of pure solvent and the solution containing the solute atoms. For dilute solutions, one can define a linear regression coefficient defined as  $k = N_{\text{sol}} \Delta a$  with  $N_{\text{sol}}$  being the number of atoms in the supercell. The linear regression coefficients and local relaxations for the ten elements listed have been evaluated from first-principles calculations.<sup>44</sup>

Interfacial energies between  $\gamma$  and  $\gamma'$  have been calculated for two types of interfaces, (001) and (011), with four cubic L1<sub>2</sub>  $\gamma'$ -Ni<sub>3</sub>Al unit cells and four cubic fcc  $\gamma$ -Ni cells. The Monkhorst k points of  $12 \times 12 \times 6$  and  $8 \times 12 \times 8$  are employed for (001) and (011) interfaces, respectively. The interfacial energies thus obtained are 39.6 mJ/m<sup>2</sup> and 63.8 mJ/m<sup>2</sup> for (100) and (110), respectively.

In the Ni-Al-Mo ternary system, there are three binary systems: Ni-Al, Al-Mo, and Ni-Mo. The thermodynamic models of Ni-Al<sup>48</sup> and Al-Mo<sup>49</sup> were taken from the literature. In the Ni-Mo system, the stability of Ni<sub>2</sub>Mo, Ni<sub>3</sub>Mo, and  $\delta$ -NiMo was recently investigated by first-principles calculations.<sup>50</sup> Based on those newly obtained enthalpy of formation of compounds, the enthalpy of mixing in fcc and body-centered cubic by SQS, and experimental data in the literature, the thermodynamic modeling of the Ni-

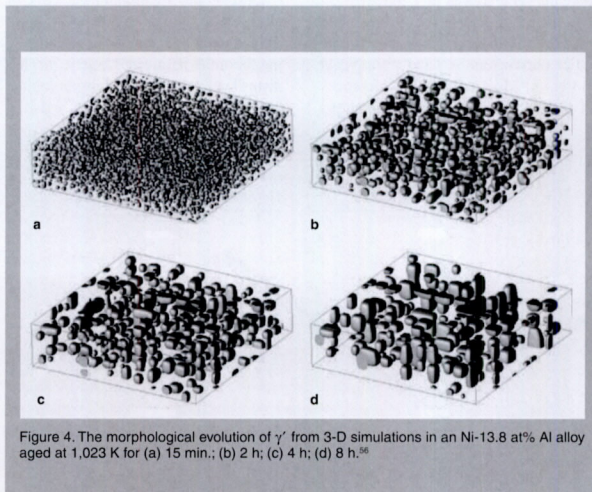


Figure 4. The morphological evolution of  $\gamma'$  from 3-D simulations in a Ni-13.8 at% Al alloy aged at 1,023 K for (a) 15 min.; (b) 2 h; (c) 4 h; (d) 8 h.<sup>36</sup>



Mo system was carried out and a new Ni-Mo phase diagram was obtained.<sup>51</sup> In the Ni-Al-Mo system, the parameters of the non-stoichiometric phases from three binary systems were evaluated by considering the solubility of the third component, and ternary phases introduced based on first-principles calculations.<sup>52</sup>

The atomic mobility in  $\gamma$  Ni-Al was modeled by A. Engstrom and J. Ågren.<sup>53</sup> The atomic mobility in  $\gamma'$  of Ni-Al is needed along with the atomic mobility in  $\gamma$  and  $\gamma'$  of Ni-Mo, Al-Mo, and Ni-Al-Mo. Based on available experimental information, the atomic mobility for Al-Mo and Ni-Mo were evaluated.<sup>54</sup> For atomic mobility in  $\gamma'$ , to extract the ordering effect from the experimental data, both the kinetic description for the related disordered phase and the related thermodynamic description are needed.

The thermodynamic database for Ni-Al from N. Dupin et al.<sup>48</sup> and the mobility database for  $\gamma$  Al-Ni from Engstrom and Ågren<sup>53</sup> were used to evaluate the model parameters in  $\gamma'$ . Since molybdenum prefers to occupy aluminum sites in  $\gamma'$ , the effect of Al-Mo ordering is ignored and the atomic mobility of molybdenum in  $\gamma'$ -Ni<sub>3</sub>Al is assumed to be the same as that of aluminum. The atomic mobility in the hypothetical  $\gamma'$ -Ni<sub>3</sub>Mo phase is assumed to be the same as that in  $\gamma'$ -Ni<sub>3</sub>Al.<sup>54</sup>

Lattice parameters in Ni-Al have been modeled.<sup>54</sup> Model parameters for the lattice parameters of  $\gamma$  and  $\gamma'$  in Ni-Al and Ni-Al-Mo are listed in References 34 and 54. Using the lattice parameters

Equations	
$G_m = {}^{ref}G_m + {}^{ideal}G_m + {}^{ss}G_m$	(1)
$D_i = RTM_i$	(2)
$M_i = \frac{D_i^*}{RT} = \frac{D_i^*}{RT} \exp\left(-\frac{Q_i}{RT}\right)$ $= \frac{1}{RT} \exp\left[-\frac{1}{RT}(Q_i - RT \ln D_i^*)\right]$ $= \frac{1}{RT} \exp\left[-\frac{\Delta\Phi_i}{RT}\right]$	(3)
Bulk modulus = $\frac{1}{4} \cdot 1.00096 \text{ EN} - 0.35682 \text{ x} - 0.77228 \text{ BLA-N}$ $- 0.83367 \text{ BLB-Np} \cdot 0.03296 \text{ Q}^* \text{ tet} \cdot 0.18484 \text{ Q}^* \text{ oct} - 0.13503 \text{ Q}^* \text{ N}$	(4)

of  $\gamma$  and  $\gamma'$ , the lattice mismatch between them was predicted. Extreme care must be taken in using proper phase compositions based on experimental procedures.<sup>54</sup> The effect of molybdenum on the lattice parameters of  $\gamma$  and  $\gamma'$  was studied through first-principles calculations with its substitution in either sublattice considered separately and modeled through the lattice change  $\Delta a_c = \sum \sum k_i y_i^s$  where  $s$  indicates different sublattices and  $i$  denotes solute or anti-site elements,  $y_i^s$  is the atomic fraction of  $i$  in sublattice  $s$ , and  $k_i^s$  is the related linear coefficient. Combining with the lattice parameter of  $\gamma'$ -Ni<sub>3</sub>Al, the lattice mismatch between  $\gamma$  and  $\gamma'$  is calculated.<sup>44,54</sup>

Despite the remarkable success of phase-field simulations in providing a fundamental understanding of underlying thermodynamic and kinetic mechanisms leading to various morphological evolutions of  $\gamma'$  precipitates, existing

phase-field simulation results are largely qualitative due to the simplifications in thermodynamic and kinetic properties and the existing numerical approaches and computer power that limit most simulations to two dimensions. However, there has been significant recent progress in developing phase-field models for predicting the quantitative coarsening kinetics of  $\gamma'$  precipitates.<sup>55-57</sup> In these simulations, nucleation of  $\gamma'$  in the supersaturated  $\gamma$  phase was explicitly introduced into the system at the beginning of a phase-field simulation using an approach similar to that described by Simmons et al.<sup>58</sup>

Figure 4 shows an example of three-dimensional (3-D) morphological evolution of  $\gamma'$  in an Ni-13.8 at.% Al alloy at a temperature of 1,023 K. The simulation was performed using a  $64 \times 256 \times 256$  grid with a unit grid size of 2.5 nm. In other words, the system is  $160 \text{ nm} \times 640 \text{ nm} \times 640 \text{ nm}$  in its physical size. The initial state was a homogeneous solution with small composition fluctuations around the average composition. Spherical nuclei with an average radius of 7.5 nm were randomly introduced to the computational domain at  $t > 0$  with the mass balance maintained at all times. Approximately 6,000 particles were formed during the nucleation stage. According to experimental observations,<sup>59</sup> precipitation occurs very rapidly in this system. Thus, the nucleation event was switched off at an early stage of aging ( $t = 5 \text{ min}$ ).

The volume fraction of  $\gamma'$  at the four different aging times in Figure 4 is approximately 9–12%, which is very close to the equilibrium volume fraction

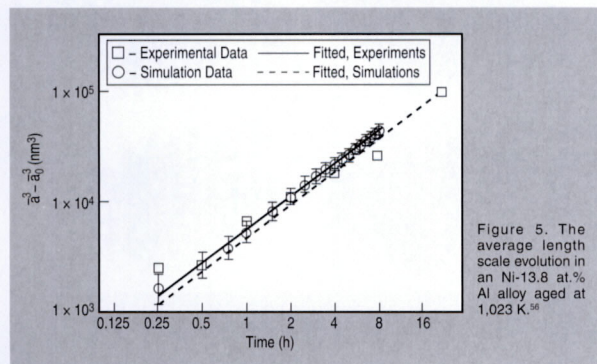
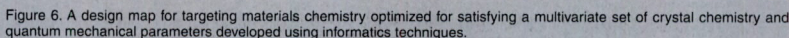


Figure 5. The average length scale evolution in an Ni-13.8 at.% Al alloy aged at 1,023 K.<sup>56</sup>

With the precipitate morphologies as a function of time, one can evaluate the quantitative coarsening kinetics. For cube-shaped particles, the relationship between the average cube edge length  $\bar{a}$  and aging time can be expressed as  $\bar{a}^3 - \bar{a}_0^3 = Kt$ . The mean particle edge length was obtained by following the

Recently, phase-field simulations of the morphological evolution and coarsening kinetics in Ni-Al-Mo have also been carried out.<sup>54</sup> To compare with

It is observed that the morphologies of  $\gamma'$  in the simulated and experimental microstructures agree well with each other. With an increase of the molybdenum concentration from A1 to A3, the lattice mismatch decreases, and the  $\gamma'$  morphology changes from rectangular to circular. For all the alloy compositions, the coalescence is observed between neighboring domains with the same order parameter, similar to the previous phase-field simulations in Ni-Al alloys.<sup>56</sup> The coarsening rate constants as a function of molybdenum contents and  $\gamma'$  volume fractions have been compared with experimental data in the literature.<sup>54</sup>





## SOFT MODELING: APPLICATION TO CRYSTAL CHEMISTRY DESIGN

As an example of the soft modeling approach to "materials-by-design," the authors have worked to identify combinations of mixed metal sequences for a given stoichiometry of nitrides that can serve as promising candidates for the desired properties.<sup>61</sup> To experimentally synthesize and characterize a vast array of possible compositions is, of course, not realistic. Similarly, despite the advances and sophistication of first-principles predictions, it is equally prohibitive to computationally predict properties for a large combinatorial array of crystal chemistries. The informatics approach is unique in that it blends in data derived from both experiments and computation to serve as a "training" data set and integrates it into an ultra-large elemental database. This database contains a vast array of property information that is relevant to issues of interest here such as thermal expansion, compressibility, and modulus.

The fundamental strategy of this work is to apply multivariate analysis on the multiple parameters or latent variables (LVs), which are used in the theoretical calculations to quantitatively assess the statistical relationship of each of the numerous descriptors of each compound studied.<sup>62-74</sup> The goal is to explore the individual correlations of the specific variables (i.e., LVs) used in the ab-initio calculations, which take into account their relative impact on final properties. The partial least squares (PLS) regression method is particularly appropriate for QSAR formulations as it is used to predict properties based on variables (even some which may have only indirect impact) which collectively relate to these properties. In addition, model parameters in PLS can be more accurately calculated with an increasing number of relevant variables and observations. Partial least squares also has the advantage over multiple linear regressions for handling of collinearity and missing data. The mathematical formulation of this method is one that calculates a projection which captures the variance in the predictor variables (X) and the correlation between it and the predicted variables (Y). In other words, PLS is a method to maxi-

mize covariance between factors, X, and responses, Y. Detailed geometrical or mathematical descriptions can be found in the literature.<sup>75</sup> The PLS analysis results in an equation that is a linear combination of the LVs.

The emphasis in PLS is on predicting the responses and not necessarily on trying to understand the underlying relationship between the variables (hence the term "soft modeling"). For example, PLS is not usually appropriate for screening out factors that have a negligible effect on the response. However, when prediction is the goal and there is no practical need to limit the number of measured factors, PLS can be a useful

**The PLS method  
can be applied to  
rationalize the  
materials attributes  
relevant to materials  
function or property.**

tool. If the number of factors gets too large (for example, greater than the number of observations as in the case of multi-scale modeling), one is likely to get a model that fits the sampled data perfectly but that will fail to predict new data well. This phenomenon is called over-fitting. In such cases, although there are many manifest factors, there may be only a few underlying or latent factors that account for most of the variation in the response. The general idea of PLS is to try to extract these latent factors, accounting for as much of the manifest factor variation as possible while modeling the responses well. For this reason, the acronym PLS has also been taken to mean projection to latent structure. The PLS method can be applied to rationalize the materials attributes relevant to materials function or property. This permits one to use PLS methods to develop explicit quantitative relationships that identify the relative contributions of different data descriptors, and the resulting relationship between all these descriptors as a linear combination, to the final property.

For instance C. Suh and K. Rajan<sup>61</sup> explored the attributes used in electronic structure calculations and their influence

on predicting bulk modulus. Using PLS, a QSAR was developed relating bulk modulus with a variety of

$$\begin{aligned} \text{Bulk modulus} = & \frac{1}{4} \cdot 1.00096 \text{ EN} \\ & - 0.35682 x - 0.77228 \text{BLA-N} \\ & - 0.83367 \text{BLB-N} p 0.03296 Q^{\text{tet}} p \\ & 0.18484 Q^{\text{oct}} - 0.13503 Q^{\text{N}} \end{aligned}$$

where EN = weighted electronegativity difference,  $x$  = internal anion parameter; BLA-N = A-N bond length; BLB-N = B-N bond length;  $Q^{\text{tet}}$  = Mulliken effective charge for tetrahedral site ion;  $Q^{\text{oct}}$  = Mulliken effective charge for octahedral site ion; and  $Q^{\text{N}}$  = Mulliken effective charge for N ion.

By systematically exploring the number and type of variables needed, very strong agreement was found in being able to predict properties consistent with ab-initio calculations based strictly on a data-driven analysis. Based on the authors' QSAR formulation, the role of the effective charge ( $Q^*$ ) in enhancing modulus is particularly notable. This is consistent with theoretical studies that show it is the effective charge parameter which helps to define the degree of charge transfer and the level of covalency associated with the specific site occupancy of a given species. Ab-initio calculations of this effective charge can then be used as a major screening parameter in identifying promising crystal chemistries for promoting the modulus. Hence, using PLS to develop a QSAR formulation combined with an interpretation of the physics governing these materials can indeed be valuable. The authors' predictions fit well with systems of similar electronic structure and allow the identification of outliers based on these quantum mechanical calculations. Based on these predictions, materials design can be effectively accelerated by focusing on promising candidate chemistries. Those selected can then be subjected to further analysis via experimentation and computational methods to validate crystal-structure-level properties. The data generated by these selective experiments and computations also serve to refine the next generation of training data for another iterative round of data mining, which permits a further refinement of high-throughput predictions.

By using a "rational drug design" approach to targeting potential new crystal chemistries, Suh and Rajan



have developed virtual compound chemistries with targeted modulus properties. This approach can be extended to cover a broad range of refractory metal alloys, oxides, nitrides, and oxy-nitrides, and will be complemented by ab-initio calculations to study interactions with these materials. The informatics will help to facilitate the multiscale modeling as shown in the clear delineation of the modulus behavior associated with refractory materials from oxides and nitrides and diamond. When coupled to processing and performance data, models can be built for establishing structure-property relationships to serve as design maps for targeting materials chemistry optimized for satisfying a multivariate set of design parameters, as illustrated in Figure 6.

## CONCLUSION

The two computational strategies for linking length scales that were presented in this report can help a materials-by-design strategy. The integration of both these approaches produces another tool for materials scientists to accelerate design and discovery.

## ACKNOWLEDGEMENTS

Zi-Kui Liu and Long-Qing Chen are grateful for the financial support from NASA under grant no. NCC3-920 and the National Science Foundation (NSF) under grants DMR-9983532, DMR-0122638, and DMR-0205232. Many postdoctoral fellows and graduate students in the authors' research labs at Pennsylvania State University contributed to the research activities, particularly Dr. Raymundo Arroyave, Dr. Chao Jiang, Dr. Yi Wang, Dr. Shihui Zhou, Dr. Jingzhi Zhu, and Dr. Tao Wang. The authors greatly appreciate the collaborations with Dr. Chris Wolverton at Ford Motor Company and Dr. Jorge Sofo at Pennsylvania State University. Krishna Rajan is grateful for financial support from the NSF International Materials Institute program for Combinatorial Sciences and Materials Informatics Collaboratory DMR-06-03644; the National Science Digital Library—Materials Digital Library Pathway project: DUE-05-323831; and the Air Force Office of Scientific Research—Informatics for Alloy Design project. He is particularly grateful for the contributions from his group, especially Dr. C. Suh.

## References

1. Z.-K. Liu et al., "An Integrated Framework for Multi-Scale Materials Simulation and Design," *J. Comput.-Aided Mater. Des.*, 11 (2004), pp. 183-199.
2. R. Arroyave, A. van de Walle, and Z.-K. Liu, "First-Principles Calculations of the Zn-Zr System," *Acta Mater.*, 54 (2006), pp. 473-482.
3. S. Shang et al., "Phase Stability in a- and b-Rhombohedral Boron," *Phys. Rev. Lett.*, 2006, submitted.
4. Y. Wang, Z.-K. Liu, and L.-Q. Chen, "Thermodynamic Properties of Al, Ni, NiAl, and NiAl from First-Principles Calculations," *Acta Mater.*, 52 (2004), pp. 2665-2671.
5. R. Arroyave, D. Shin, and Z.-K. Liu, "Ab initio Thermodynamic Properties of Stoichiometric Phases in the Ni-Al System," *Acta Mater.*, 53 (2005), pp. 1809-1819.
6. A. van de Walle, M. Asta, and G. Ceder, "The Alloy Theoretic Automated Toolkit: A User Guide," *CALPHAD*, 26 (2002), pp. 539-553.
7. A. E. Kissavos et al., "A Critical Test of ab initio and CALPHAD Methods: The Structural Energy Difference between bcc and hcp Molybdenum," *CALPHAD*, 29 (2005), pp. 17-23.
8. C. Wolverton and A. Zunger, "Ising-like Description of Structurally Released Ordered and Disordered Alloys," *Phys. Rev. Lett.*, 75 (1995), pp. 3162-3165.
9. V. Ozolins, C. Wolverton, and A. Zunger, "Cu-Au, Ag-Au, Cu-Ag, and Ni-Au Intermetallics: First-Principles Study of Temperature-Composition Phase Diagrams and Structures," *Phys. Rev. B*, 57 (1998), pp. 6427-6443.
10. A. Zunger et al., "Special Quasirandom Structures," *Phys. Rev. Lett.*, 65 (1990), pp. 353-356.
11. D. Shin et al., "Thermodynamic Properties of Binary hcp Solution Phases from Special Quasirandom Structures," *Phys. Rev. B*, 74 (2006), p. 024204.
12. C. Jiang et al., "First-Principles Study of Binary bcc Alloys using Special Quasirandom Structures," *Phys. Rev. B*, 69 (2004), p. 214202.
13. C. Jiang, L.-Q. Chen, and Z.-K. Liu, "First-Principles Study of Constitutional Point Defects in B2 NiAl Using Special Quasirandom Structures," *Acta Mater.*, 53 (2005), pp. 2643-2652.
14. Y. Zhong et al., "First-Principles Investigation of Laves Phases in Mg-Al-Ca System," *Mater. Sci. Forum*, 488-489 (2005), pp. 169-175.
15. M. Yang and Z.-K. Liu, "SQS for Halite," unpublished research (2005).
16. D. Shin, R. Arroyave, and Z.-K. Liu, "Thermodynamic Properties of Binary HCP Solution Phases from Special Quasirandom Structures," *Phys. Rev. B* (submitted 2005).
17. T. Wang, L.-Q. Chen, and Z.-K. Liu, "SQS for  $L_{12}$ ," unpublished research (2005).
18. M. Sluiter and Y. Kawazoe, "Prediction of Matrix-Precipitate Interfacial Free Energies: Application to Al-Al<sub>3</sub>," *Phys. Rev. B*, 54 (1996), pp. 10381-10384.
19. V. Vaityanathan, C. Wolverton, and L.-Q. Chen, "Multiscale Modeling of Precipitate Microstructure Evolution," *Phys. Rev. Lett.*, 88 (2002), p. 125503.
20. C. Colinet, "Ab initio Determination of the (001) Antiphase-Boundary Energy in the  $Do_{19}$  Ni<sub>3</sub>V Compound," *Philos. Mag. B-Phys. Condens. Matter Sci. Mech. Electron. Opt. Magn. Prop.*, 82 (2002), pp. 1715-1729.
21. N. Sandberg, B. Magyari-Kope, and T.R. Mattsson, "Self-Diffusion Rates in Al from Combined First-Principles and Model-Potential Calculations," *Phys. Rev. Lett.*, 89 (2002), p. 065901-1.
22. A. Van Der Ven and G. Ceder, "First Principles Calculation of the Interdiffusion Coefficient in Binary Alloys," *Phys. Rev. Lett.*, 94 (2005), p. 045901.
23. Y. Le Page and P. Saxe, "Symmetry-General Least-Squares Extraction of Elastic Data for Strained Materials from ab initio Calculations of Stress," *Phys. Rev. B*, 65 (2002), p. 104104.

24. L. Kaufman and H. Bernstein, *Computer Calculation of Phase Diagram* (New York: Academic Press Inc., 1970).
25. U.R. Kattner et al., "Applications of Computational Thermodynamics: Groups 4 and 5: Use of Thermodynamic Software in Process Modelling and New Applications of Thermodynamic Calculations," *CALPHAD*, 24 (2000), pp. 55-94.
26. L. Kaufman, "Computational Thermodynamics and Materials Design," *CALPHAD*, 25 (2001), pp. 141-161.
27. I. Ansara and B. Sundman, "The Scientific Group Thermodata Europe," *Computer Handling and Dissemination of Data*, ed. P.S. Glaeser (New York: Elsevier Science Pub. Co., 1987), pp. 154-158.
28. A.T. Dinsdale, "SGTE Data for Pure Elements," *CALPHAD*, 15 (1991), pp. 317-425.
29. N. Saunders and A.P. Miodownik, *CALPHAD (Calculation of Phase Diagrams): A Comprehensive Guide* (Oxford and New York: Pergamon, 1998).
30. I. Ansara et al., "Thermodynamic Assessment of the Al-Ni System," *J. Alloy. Compd.*, 247 (1997), pp. 20-30.
31. J.O. Andersson and J. Agren, "Models for Numerical Treatment of Multicomponent Diffusion in Simple Phases," *J. Appl. Phys.*, 72 (1992), pp. 1350-1355.
32. T. Helander and J. Agren, "A Phenomenological Treatment of Diffusion in Al-Fe and Al-Ni Alloys having B2-BCC Ordered Structure," *Acta Mater.*, 47 (1999), pp. 1141-1152.
33. X.G. Lu, M. Selby, and B. Sundman, "Theoretical Modeling of Molar Volume and Thermal Expansion," *Acta Mater.*, 53 (2005), pp. 2259-2272.
34. T. Wang et al., "Modeling of Lattice Parameter in the Ni-Al System," *Mater. Trans. A*, 35A (2004), pp. 2313-2321.
35. L.-Q. Chen, "Phase-Field Models for Microstructure Evolution," *Ann. Rev. Mater. Res.*, 32 (2002), pp. 113-140.
36. J.W. Cahn and J.E. Hilliard, "Free Energy of a Nonuniform System. I. Interfacial Free Energy," *J. Chem. Phys.*, 28 (1958), pp. 258-267.
37. J.W. Cahn, "On Spinodal Decomposition," *Acta Metall.*, 9 (1961), pp. 795-801.
38. S.M. Allen and J.W. Cahn, "A Microscopic Theory of Domain Wall Motion and Its Experimental Verification in Fe-Al Alloy Domain Growth Kinetics," *J. de Physique*, C7 (1977), pp. C7-51.
39. G. Cagin and W. Xie, "Phase-Field and Sharp-Interface Alloy Models," *Phys. Rev. E*, 48 (1993), pp. 1897-1909.
40. A. Karma and W.J. Rappel, "Phase-Field Method for Computationally Efficient Modeling of Solidification with Arbitrary Interface Kinetics," *Phys. Rev. E*, 53 (1996), pp. R3107-R3120.
41. S.G. Kim, W.T. Kim, and T. Suzuki, "Phase-Field Model for Binary Alloys," *Phys. Rev. E*, 60 (1999), pp. 7186-7197.
42. A. Karma, "Phase-Field Formulation for Quantitative Modeling of Alloy Solidification," *Phys. Rev. Lett.*, 87 (2001), art. no. 115701.
43. A. Baldan, "Review Progress in Ostwald Ripening Theories and Their Applications to the  $\gamma$ -Precipitates in Nickel-Base Superalloys: Part II. Nickel-Base Superalloys," *J. Mater. Sci.*, 37 (2002), p. 2379.
44. T. Wang, L.-Q. Chen, and Z.-K. Liu, "Lattice Parameters and Local Lattice Distortion in fcc-Ni Solutions," *Acta Mater.*, submitted (2005).
45. V. Biss and D.L. Sponser, "Effect of Molybdenum on Gamma Prime Coarsening and on Elevated-Temperature Hardness in Some Experimental Ni-Base Superalloys," *Mater. Trans.*, 4 (1973), pp. 1953-1960.
46. G. Kresse and J. Furthmüller, *Vienna ab initio simulation package (VASP)* (2003), <http://cms.mpi.univie.ac.at/vasp/vasp.html>.
47. J.P. Perdew and Y. Wang, "Accurate and Simple Analytic Representation of the Electron-Gas Correlation-Energy," *Phys. Rev. B*, 45 (1992), pp. 13244-13249.



48. N. Dupin, I. Ansara, and B. Sundman, "Thermodynamic Re-Assessment of the Ternary System Al-Cr-Ni," *CALPHAD*, 25 (2001), pp. 279-298.

49. N. Saunders, "The Al-Mo System (Aluminum-Molybdenum)," *J. Phase Equilib.*, 18 (1997), pp. 370-378.

50. Y. Wang et al., "Structural Stability of Ni-Mo Compounds from First-Principles Calculations," *Scr. Mater.*, 52 (2005), pp. 17-20.

51. S.H. Zhou et al., "First-Principles Calculations and Thermodynamic Modeling of the Ni-Mo System," *Mater. Sci. Eng. A*, 397 (2005), pp. 288-296.

52. S.H. Zhou et al., "Computational Tools for Designing Ni-Based Superalloys," *Superalloy 2004*, ed. K.A. Green et al. (Warrendale, PA: TMS, 2004), pp. 969-975.

53. A. Engstrom and J. Agren, "Assessment of Diffusional Mobilities in Face-Centered Cubic Ni-Cr-Al Alloys," *Z. Metallkd.*, 87 (1996), pp. 92-97.

54. T. Wang, "Integrated Computational Tools for Microstructure Simulations: Applications to Ni-Al-Mo Alloys" (Ph.D. Thesis, Pennsylvania State University, 2006).

55. J.Z. Zhu et al., "Linking Phase-Field Model to CALPHAD: Application to Precipitate Shape Evolution in Ni-Based Alloys," *Scr. Mater.*, 46 (2002), pp. 401-406.

56. J.Z. Zhu et al., "Three-Dimensional Phase-Field Simulations of Coarsening Kinetics of  $\gamma$  Particles in Binary Ni-Al Alloys," *Acta Mater.*, 52 (2004), pp. 2837-2845.

57. Y.H. Wen et al., "A Phase-Field Model for Heat Treatment Applications in Ni-Based Alloys," *Acta Mater.*, 54 (2006), pp. 2087-2099.

58. J.P. Simmons, C. Shen, and Y. Wang, "Phase Field Modeling of Simultaneous Nucleation and Growth by Explicitly Incorporating Nucleation Events," *Scr. Mater.*, 43 (2000), pp. 935-942.

59. A.J. Ardell and R. Nicholson, "Coarsening of  $\gamma$  in Ni-Al Alloys," *J. Phys. Chem. Solids*, 27 (1966), pp. 1793-1800.

60. M. Fahrman et al., "Influence of Coherency Stress on Microstructural Evolution in Model Ni-Al-Mo Alloys," *Acta Metall. Mater.*, 43 (1995), pp. 1007-1022.

61. C. Suh and K. Rajan, "Virtual Screening and QSAR Formulations for Crystal Chemistry," *QSAR & Combinatorial Science Journal*, 24 (2005), p. 114.

62. K. Rajan and A. Rajagopalan "Informatics Based Optimization of Crystallographic Descriptors for Framework Structures," *Combinatorial and High Throughput Discovery and Optimization of Catalysts and Materials*, ed. W. Maier and R.A. Potyrailo (Boca Raton, FL: CRC Press, 2006).

63. M. Stukowski et al., "Informatics for Combinatorial Experiments: Accelerating Data Interpretation," *Combinatorial Methods and Informatics in Materials Science*, ed. Q. Wang et al. (Warrendale, PA: Materials Research Society, 2006), in press.

64. K. Rajan, C. Suh, and B. Narasimhan, "Informatics Methods for Combinatorial Materials Science," *Combinatorial Materials Science*, ed. S.K. Mallapragada, B. Narasimhan, and M.D. Porter (New York: John Wiley & Sons, in press).

65. K. Rajan, "Materials Informatics," *Materials Today* (October 2005), pp. 38-45.

66. K. Rajan, "An Informatics Approach to Interface Characterization: Establishing a 'Materials by Design' Paradigm," *Science and Technology of Interfaces*, ed. S. Ankem and C.S. Pande (Warrendale, PA: TMS, 2002), pp. 231-242.

67. K. Rajan et al., "Quantitative Structure-Activity Relationships (QSARs) for Materials Science," *Combinatorial and Artificial Intelligence Methods in Materials Science*, vol. 700, ed. Ichiro Takeuchi et al. (Warrendale, PA: MRS, 2002), pp. S7.51-S7.510.

68. C. Suh et al., "Applications of Principal Component Analysis in Materials Science," *Data Science Journal*, 1 (2002), p. 19.

69. K. Rajan, A. Rajagopalan, and C. Suh, "Data Mining and Multivariate Analysis in Materials Science," *Molten Salts—Fundamentals to Applications*, ed. M. Gaune-

Escard (Norwell, MA: Kluwer Academic, 2002), pp. 241-248.

70. C. Suh et al., "Chemical Discovery in Molten Salts through Data Mining," *International Symposium on Ionic Liquids: Festschrift in honor of Prof. M. Gaune-Escard*, ed. H.A. Oye and A. Jagtveit (Trondheim, Norway: Norwegian University of Science and Technology, 2003), pp. 587-599.

71. C. Suh and K. Rajan, "Combinatorial Design of Semiconductor Chemistry for Bandgap Engineering: 'Virtual' Combinatorial Experimentation," *Applied Surface Science*, 223 (2003), p. 148.

72. A. Rajagopalan et al., "Secondary Descriptor Development for Zeolite Framework Design: An Informatics Approach," *Applied Catalysis A*, 254 (2003), pp. 147-160.

73. A. Rajagopalan et al., "An Informatics Approach to Materials Design," *Proc. 7th Intl. Conference on Systems, Cybernetics and Informatics* (Orlando, FL: Intl. Institute of Informatics and Systems, 2003).

74. C. Suh et al., "Combinatorial Materials Design Through Database Science," *Combinatorial and Artificial Intelligence Methods in Materials Science II*, vol. 804, ed. Radislav A. Potyrailo et al. (Warrendale, PA: Materials Research Society, 2004), pp. 23.1-23.8.

75. L. Eriksson et al., *Multi- and Megavariate Data Analysis—Principles and Applications* (Umeå, Sweden: Umetrics Academy, 1999).

Zi-Kui Liu and Long-Qing Chen are with the Department of Materials Science and Engineering at the Pennsylvania State University. Krishna Rajan is with the Department of Materials Science and Engineering at Iowa State University in Ames, Iowa.

For more information, contact Krishna Rajan, Iowa State University, Department of Materials Science and Engineering, 22205 Hoover Hall, Ames, IA 50011; (515) 294-2670, e-mail krajan@iastate.edu.

## RECENTLY REVIEWED BOOKS

Available Only at JOM On-Line

The following materials science and engineering publications have recently been reviewed by JOM readers. To read the reviews, or to become a JOM book reviewer, visit the JOM web site and click the Book Review link (under Complimentary On-Line Content). All book reviews are published exclusively on the JOM web site.

### Titanium Alloys: Russian Aircraft and Aerospace Applications (2006)

Valentin N. Moiseyev, author

ISBN 0-8493-3273-7. Taylor & Francis Group, Boca Raton, Florida, 2006.

Hardcover: 206 pages. \$159.95.

Reviewed by: Mary Raum, Naval War College

### Inorganic Materials Chemistry Desk Reference, Second Edition (2005)

D. Sangeeta and John R. LaGraff, authors

ISBN 0-8493-0910-7. CRC Press, Boca Raton, Florida, 2005. Hardcover.

372 pages. \$159.95.

Reviewed by: Maria Salamon, Naval War College

### Introduction to Materials Modelling (2005)

Zoe H. Barber, editor

ISBN 1-902653-76-9. Maney Publishing for the Institute of Materials, Minerals, and Mining, London, 2005. Softcover. 165 pages. \$31.00.

Reviewed by: Patrice E.A. Turchi, Lawrence Livermore National Laboratory

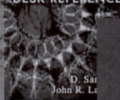
### Titanium Alloys

Russian Aircraft and Aerospace Applications



Valentin N. Moiseyev

### SECOND EDITION INORGANIC MATERIALS CHEMISTRY DESK REFERENCE



### Introduction to Materials Modelling

Edited by Zoe H. Barber



Read them on-line at:  
[www.tms.org/jom.html](http://www.tms.org/jom.html)  
Click on Book Reviews