**Supplementary Material**

Correlation analysis of properties by machine learning: Illustrated with stacking fault energy from first-principles calculations in dilute fcc-based alloys

Xiaoyu Chong,1,2 Shun-Li Shang,2,\* Adam M. Krajewski,2 John D. Shimanek,2 Weihang Du,2 Yi Wang,2 Jing Feng,1,\* Dongwon Shin,3 Allison M. Beese,2 and Zi-Kui Liu2

1Faculty of Material Science and Engineering, Kunming University of Science and Technology, Kunming 650093, China

2Department of Materials Science and Engineering, The Pennsylvania State University, University Park, PA, 16802, USA

3Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA

\*Emails: sus26@psu.edu (S. L. Shang) and jingfeng@kust.edu.cn (J. Feng)

Table S 1. Elemental descriptors examined in the present work together with their |PCC| and MIC values between the descriptors and the DgSFE values. The numbers in “Rank” column rank the statistical significances of the descriptors to regulate DgSFE after iterations between sequential descriptor selection and ML predictions using the GRP4 model (see details in Table 3). The values of these descriptors are given in the supplemental csv (comma-separated value) file, Original\_Descriptors49.csv.

Str1: Structure (BCC, FCC, HCP)

USFE: Unstable stacking fault energy (J/m2)

ISIS: ideal shear strength (GPa)

SurfEne: Surface energy (J/m2)

WorkFunc: Work function

PackingEff

DFTBhvGh: B/G

DFTGvb: G/b

DFTGV: G\*V

ratio\_bh: b/h

Gb2pih: G\*b/(2ph)

BGsq: (B/G)^(1/2)

GGB32: (G^3/B^2)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Rank | Descriptors |  | MIC | Explanations |
| 1 | Radius\_Coval | 0.790 | 0.716 | Covalent radius (in pm) based on the collections of Wolfram Mathematica (WM); see “ElementData” in WM [1]. |
| 2 | NpVal | 0.05 | 0.373 | Number of p valence electrons. |
| 3 | Group | 0.438 | 0.648 | Group of pure elements in the periodic table. |
| 4 | NdVal | 0.406 | 0.319 | Number of d valence electrons. |
| 5 | Heat\_Sublimation | 0.389 | 0.406 | Heat of sublimation at 298 K (in J/mol) [2]. |
| 6 | G\_wiki | 0.231 | 0.310 | Shear modulus (in GPa) of pure elements based on Wikipedia and  Azom [3,4]. |
| 7 | Nval | 0.068 | 0.315 | Number of total valence electrons. |
| 8 | Va\_Form\_FCC | 0.102 | 0.331 | Predicted vacancy formation energies of pure elements in fcc structure [5]. |
| 9 | MeltingT | 0.254 | 0.437 | Melting temperature (in K) based on the collections by Kittel [6]. |
| 10 | EleDensity\_Miedema | 0.400 | 0.357 | Electron density at the boundary of Wigner-Seitz cell used in the Miedema model[7]. |
| 11 | EleNeg\_Miedema | 0.600 | 0.405 | Electronegativity (in Volt) used in the Miedema model[7]. |
| 12 | Period | 0.290 | 0.326 | Period of pure elements in the periodic table. |
| 13 | S298 | 0.379 | 0.354 | Standard entropy at 298 K (in J/mol.K) [8]. |
| 14 | Number | 0.200 | 0.433 | Atomic number of pure elements in the periodic table. |
| 15 | PP\_radius\_s | 0.496 | 0.416 | Nonlocal pseudopotential radius for the s orbital (in a.u.) [9]. |
| 16 | NUnfill | 0.524 | 0.515 | Number of total unfilled valence electrons. |
| 17 | V0\_Miedema | 0.651 | 0.510 | Atomic volume (in cm3/mol) used in the Miedema model[7]. |
| 18 | NdUnfill | 0.524 | 0.515 | Number of unfilled d valence electrons. |
| 19 | Mass | 0.204 | 0.478 | Mass of pure elements. |
| 20 | Radius\_vDW | 0.535 | 0.449 | Van der Waals atomic radius (in pm) [1,10]. |
| 21 | Therm\_Conduc | 0.246 | 0.429 | Thermal conductivity at 300 K (in W cm-1 K-1) [6,11]. |
| 22 | Heat\_Capacity | 0.229 | 0.322 | Heat capacity at 298 K (in J/kg-mol.K) [2]. |
| 23 | Ele\_Conduc | 0.226 | 0.456 | Electrical conductivity of metals in (ohm-cm)-1 [6]. |
| 24 | B\_wiki | 0.191 | 0.275 | Bulk modulus (in GPa) of pure elements based on Wikipedia and Azom [3,4]. Note that elastic properties of fcc Sr were taken from [12]. |
| 25 | EleNeg\_Pauling | 0.398 | 0.315 | Electronegativity by Pauling scale (a dimensionless quantity) [1,10]. |
| 26 | NpUnfill | 0.059 | 0.331 | Number of unfilled p valence electrons. |
| 27 | Va\_Acti\_FCC | 0.038 | 0.309 | Predicted vacancy activity energies of pure elements in fcc structure, and the vacancy formation energies were adopted for the unstable fcc structure (i.e., Ge and La) [5]. |
| 28 | Y\_wiki | 0.219 | 0.370 | Young’s modulus (in GPa) of pure elements based on Wikipedia and Azom [3,4]. |
| 29 | Electron\_Affinity | 0.147 | 0.213 | Electron affinity (in eV) [10]. |
| 30 | NfVal | 0.162 | 0.262 | Number of f valence electrons. |
| 31 | NsVal | 0.078 | 0.215 | Number of s valence electrons. |
| 32 | PP\_radius\_p | 0.498 | 0.472 | Nonlocal pseudopotential radius for the p orbital (in a.u.) [9]. |
| 33 | Ion\_Pot\_1 | 0.409 | 0.368 | The first ionization potentials to remove one electron (in eV) [2]. |
| 34 | BoilingT | 0.260 | 0.372 | Boiling temperature (in K) [2]. |
| 35 | M\_Num2 | 0.417 | 0.703 | Mendeleev number (MN2, start from left bottom, down-top sequence) [9]. |
| 36 | NsUnfill | 0.034 | 0.237 | Number of unfilled s valence electrons. |
| 37 | MaxR\_Ele\_in\_Solid | 0.091 | 0.257 | Maximum range of electrons in solid elements for electron energy of 15 keV (in mm) [2]. |
| 38 | CohEnergy | 0.327 | 0.361 | Cohesive energy (in eV/atom) collected by Kittel [6]. |
| 39 | Heat\_Fusion | 0.255 | 0.425 | Heat of fusion at 298 K (in J/mol) [2]. |
| 40 | SpaceGroup | 0.086 | 0.285 | Space group of alloying element X. |
| 41 | Ion\_Pot\_2 | 0.206 | 0.532 | The second ionization potentials to remove two electrons (in eV) [2]. |
| 42 | Ion\_Pot\_3 | 0.322 | 0.474 | The third ionization potentials to remove three electrons (in eV) [2]. |
| 43 | DebyeT | 0.371 | 0.310 | Debye temperature (in K) collected by Kittel [6]. |
| 44 | No\_Spectral\_lines | 0.301 | 0.312 | Number of spectral lines of the elements [2]. |
| 45 | VaporHeat | 0.376 | 0.462 | Vaporization heat (in kJ/mol) based on the collections of Wolfram Mathematica (WM); see “VaporizationHeat” in WM [1]. |
| 46 | NfUnfill | 0 | 0 | Number of unfilled f valence electrons. |
| 47 | Al | 0.113 | 0.318 | Host element indicated by 1 otherwise by 0. |
| 48 | Ni | 0.161 | 0.300 | Host element Ni indicated by 1 otherwise by 0. |
| 49 | Pt | 0.045 | 0.242 | Host element indicated by 1 otherwise by 0. |

**Table S 2.** Explanation of the supplemental file, ML\_Predictions.csv.

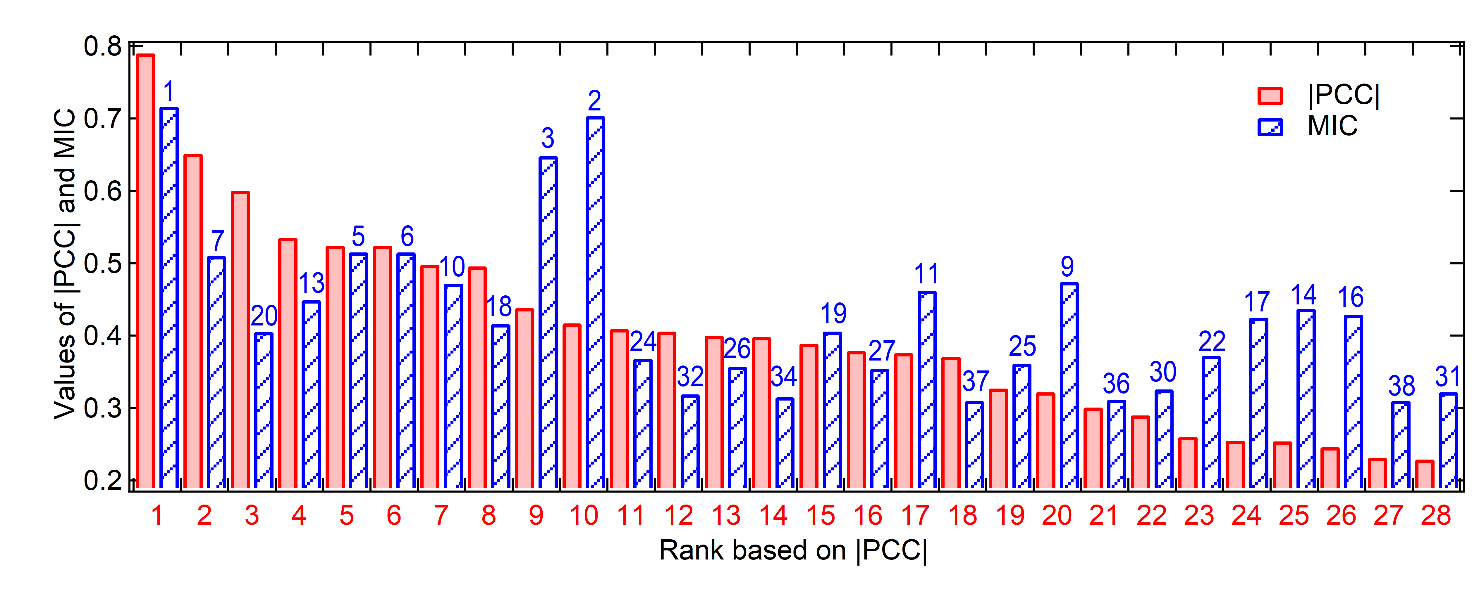
|  |  |
| --- | --- |
| Column | Note |
| Host | Host element (Al, Ni, or Pt). |
| Element | Alloying element X in Al23X, Ni23X, and Pt23X. |
| Total\_SFE | The values of stacking fault energy, gSFE,in mJ/m2. “null” indicates the value is not available based on DFT calculations. |
| Delta\_SFE | The values of relative stacking fault energy, DgSFE, in mJ/m2. “null” indicates the value is not available based on DFT calculations. See the definition of Delta\_SFE (DgSFE) in Eq. 3. |
| Pred\_Average | ML predicted DgSFE values, which are the average values in columns of Pred\_GPR2, Pred\_GPR4, Pred\_OptGPR, and Pred\_NNet. |
| Pred\_StandardDeviation | Standard deviation for values in columns of Pred\_OptGPR and Pred\_NNet. |
| Pred\_GPR2 | ML predicted DgSFE values based on the ML model by GPR2. |
| Pred\_GPR4 | ML predicted DgSFE values based on the ML model by GPR4. |
| Pred\_OptGPR | ML predicted DgSFE values based on the ML model by OptGPR. |
| Pred\_NNet | ML predicted DgSFE values based on the ML model by WM\_NNet. |
| Colum K to X | Relative values of the descriptors; see explanations for the files ReDes14.csv or ReDes49.csv (Table S4). |

**Table S 3.** Explanation of the supplemental file, Original\_Descriptors49.csv.

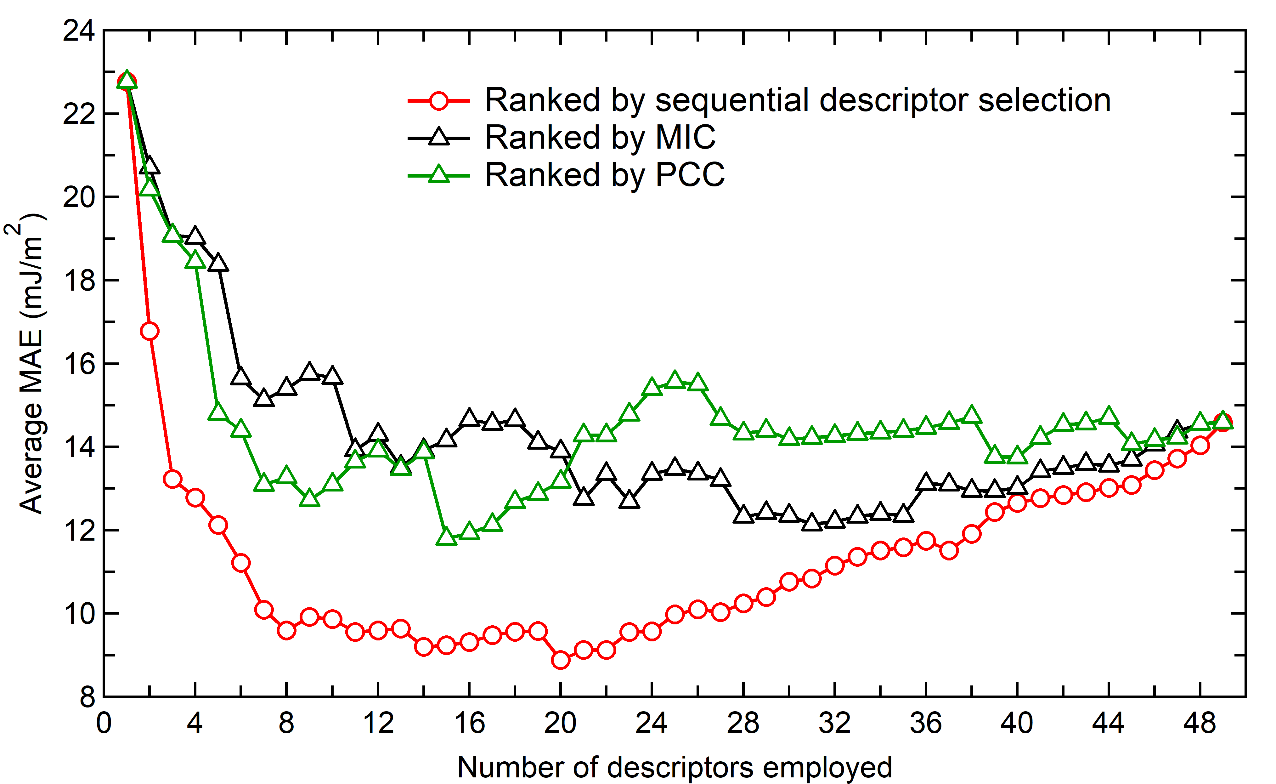
|  |  |
| --- | --- |
| Column | Note |
| Host | See explanation in Table S1. |
| Element | See explanation in Table S1. |
| Total\_SFE | See explanation in Table S1. |
| Delta\_SFE | See explanation in Table S1. |
| Radius\_vDW | Values of descriptors of the van der Waals atomic radius (in pm); see Table S1 for details. |
| … | See Table S1 for the explanations of the rest descriptors. |

**Table S 4.** Explanation of the supplemental files, ReDes49.csv and ReDes14.csv; see the definition of ReDes in Eq. 2.

|  |  |
| --- | --- |
| Column | Note |
| Host | See explanation in Table S1. |
| Element | See explanation in Table S1. |
| Total\_SFE | See explanation in Table S1. |
| Delta\_SFE | See explanation in Table S1. |
| Radius\_vDW | Relative values for the descriptor of the van der Waals atomic radius (in pm); see Eq. 2. |
| … | See Table S1 for the explanations of the rest descriptors. |

****

**Figure S 1.** Correlation analyses by |PCC| and MIC between the elemental descriptors (see the definition in Eq. 2, the explanations in Table S1, and the values in the supplemental Excel file) and the DFT-predicted DgSFE values (see the values in Table 2 and the supplemental Excel file). The integer numbers indicate the ranks for both |PCC| and MIC (see their values in Table S1).



**Figure S 2.** Average MAE values after training of 1000 times using the ReDes49 descriptors, the GPR4 algorithm, and the 5-fold cross validation method. Here, we sequentially add the descriptors for the GPR4 training according to (i) their statistic significances after sequential descriptor selections (i.e., the first generation of the ranked descriptors as detailed in Table 3), (ii) the MIC values from high to low (see Table S1), and (3) the |PCC| values from high to low (see Table S1).

**References**

[1] Wolfram Mathematica: Modern Technical Computing. Available at: https://www.wolfram.com/mathematica (Accessed: 16th March 2020).

[2] Samsonov G V. 1968 *Handbook of the Physicochemical Properties of the Elements* (New York: Springer)

[3] Elastic properties of the elements (data page) - Wikipedia. Available at: https://en.wikipedia.org/wiki/Elastic\_properties\_of\_the\_elements\_(data\_page) (Accessed: 16th March 2020).

[4] AZOM Materials. Available at: https://www.azom.com/ (Accessed: 16th March 2020).

[5] Shang S-L, Zhou B-C, Wang W Y, Ross A J, Liu X L, Hu Y-J, Fang H-Z, Wang Y and Liu Z-K 2016 A comprehensive first-principles study of pure elements: Vacancy formation and migration energies and self-diffusion coefficients *Acta Mater.* **109** 128–41

[6] Kittel C 2005 *Introduction to Solid State Physics* (Hoboken, NJ: John Wiley & Sons, Inc.)

[7] Zhang R F, Zhang S H, He Z J, Jing J and Sheng S H 2016 Miedema Calculator: A thermodynamic platform for predicting formation enthalpies of alloys within framework of Miedema’s Theory *Comput. Phys. Commun.* **209** 58–69

[8] Dinsdale A T 1991 SGTE data for pure elements *Calphad* **15** 317–425

[9] Zunger A 1980 Systematization of the stable crystal structure of all AB-type binary compounds: A pseudopotential orbital-radii approach *Phys. Rev. B* **22** 5839–72

[10] Periodic Table of Elements - PubChem. Available at: https://pubchem.ncbi.nlm.nih.gov/periodic-table/ (Accessed: 16th March 2020).

[11] The Periodic Table. Available at: https://periodictable.com (Accessed: 16th March 2020)

[12] Anderson M S, Swenson C A and Peterson D T 1990 Experimental equations of state for calcium, strontium, and barium metals to 20 kbar from 4 to 295 K *Phys. Rev. B* **41** 3329–38