Prediction System for Solid Solubility Limits of Ag-, Cu-, Al-, and Mg-Based Alloys Using Artificial Neural Networks and First-Principles Calculations

Takafumi Mochizuki^{1,*1}, Tokuteru Uesugi^{2,*2} and Yorinobu Takigawa¹

¹Department of Materials Science, Graduate School of Engineering, Osaka Prefecture University, Sakai 599-8531, Japan

²Graduate School of Humanities and Sustainable System Sciences, Osaka Prefecture University, Sakai 599-8531, Japan

A solid solubility prediction system using Hume-Rothery parameters and first-principles calculation to obtain explanatory variables was devised, and the resulting coefficients of determination, R^2 , were compared. When we used the Hume-Rothery parameter, R^2 was 0.715, and when we used the first-principles calculation results, R^2 was 0.900, indicating the improved accuracy of prediction. We tested 10-fold cross validation to evaluate the generalization performance of the network. The number of explanatory variables was optimized using the stepwise method. R^2 was maximized when eight explanatory variables were used. As a result of 10-fold cross-validation, R^2 of the constructed solid solubility prediction system which uses eight explanatory variables was 0.6993. The mean absolute error for this network was 0.45. The common logarithm value was used as the explained variable. Thus, the solid solubility limit predicted from this network was on an average 0.35 to 2.85 times the true value. [doi:10.2320/matertrans.MT-MBW2019010]

(Received July 6, 2020; Accepted July 31, 2020; Published September 18, 2020)

Keywords: first principles, artificial neural network, solid solubility limit, machine learning, Hume-Rothery

1. Introduction

The solid solubility limit plays an important role in the design of various alloys. Even if the amount of solid solution is very small, the element of the solid solution often has a profound influence on the mechanical properties of the alloy.¹⁾ Thus, the solid solubility limit has long been studied experimentally and computationally.²⁾

Hume-Rothery rules are well known semi-empirical rules to determine whether the solid metal will dissolve.³⁾ By analyzing the solid solubility limit data of Ag-based and Cu-based binary alloys, two Hume-Rothery rules were proposed. First, when the ratio of the atomic radii between the solute element and the solvent element exceeds 15%, it is difficult to form a solid solution. Second, when the ratio of the atomic radii between the solute element and the solvent element is 15% or less, the solid solution can be formed, and the range is determined by the average number of valence electrons e/a , where e is the total number of valence electrons and a is the total number of atoms. Furthermore, three rules of thumb, including the effect of electronegativity on the range of solid solution formation, were published in 1940.⁴⁾ They are known as the Hume-Rothery rules.⁵⁾

Artificial neural networks (ANN) are inspired by the human brain system.⁶⁾ ANNs are supervised machine learning algorithms that can capture the relationship between explanatory and explained variables while training a network. Recently, the interest in applying machine learning to solve problems in materials sciences has increased.^{7–13)}

First Zhang *et al.* attempted to predict the solid solubility limit for the first time using parameters of the Hume-Rothery rules and artificial neural networks.⁷⁾ Zhang *et al.* developed an artificial neural network that uses atomic size, valence, and electronegativity as explanatory variables and the solid solubility limit as an explained variable to predict the solid

solubility limit of Ag-based and Cu-based binary alloys.⁷⁾ However, the coefficient of determination, R^2 , between experimental data and predicted data obtained by the network of Zhang *et al.* was as low as 0.590.⁷⁾ Moreover, the target of the network of Zhang *et al.* was only Ag-based and Cu-based alloys.

In this study, we aimed to improve the prediction accuracy of solid solution solubility using the first-principles calculation results as explanatory variables. Moreover, we targeted not only Ag-based and Cu-based alloys to compare the accuracy with the previous result by Zhang *et al.*,⁷⁾ but also Al-based and Mg-based alloys which are widely used in various applications because of their good mechanical properties, high strength-to-weight ratio, and market acceptability. In this study, unless otherwise specified, the maximum solid solubility limit is treated as the solid solubility limit.

2. Calculation Procedures

The first-principles calculations were performed using the Cambridge Serial Total-Energy Package (CASTEP).¹⁴⁾ CASTEP is an *ab initio* pseudopotential method code for solving the electronic ground state of periodic systems, in which the wave functions are expanded on a plane-wave basis that is set using a technique based on density functional theory (DFT).^{15,16)} The electronic exchange-correlation energy used in the DFT was given by the generalized gradient approximation (GGA) proposed by Perdew *et al.* pseudopotentials¹⁷⁾ and was used for all elements.

For the calculation of Ag-, Cu-, and Al-based solid-solution alloys, we used fcc supercells that replace 1 out of 32 solvent atoms. For calculation of Mg-based solid-solution alloys, we used hcp supercells that replace 1 out of 36 solvent atoms. Figure 1 shows these supercell models. The plane-wave cut-off energy was set at 350 eV for calculations of Ag-, Al-, and Mg-based alloys and at 400 eV for calculations of Cu-based alloys. The energy integration over a Brillouin zone was

*1Graduate Student, Osaka Prefecture University

*2Corresponding author, E-mail: uesugi@kis.osakafu-u.ac.jp

Fig. 1 Schematic representation of supercell modeling (a) Ag_{31}X , Cu_{31}X , and Al_{31}X , and (b) Mg_{35}X .

performed using k -points in accordance with the Monkhorst–Pack grid¹⁸⁾ with sets of $6 \times 6 \times 6$ in calculations of Ag-based and Cu-based alloys, $7 \times 7 \times 7$ in calculations of Al-based alloys, and $6 \times 6 \times 5$ in calculations of Mg-based alloys. The spin polarizations were calculated for alloys containing Cr, Mn, Fe, Co, and Ni solute atoms which exhibit the possibility of magnetism. Full structural relaxation of the atomic configurations and lattice constants were calculated because introducing a substitute solute atom leads to local lattice distortion and changes in the cell volume. Stable atomic configurations were obtained through relaxation based on the Hellmann–Feynman forces. The lattice constants at zero pressure were also optimized using a Broyden–Fletcher–Goldfarb–Shanno (BFGS)¹⁹⁾ minimization algorithm in conjunction with the stress calculations. The convergence parameters were as follows: the total energy tolerance was 1.6×10^{-24} J/atom, the maximum force tolerance was 4.8×10^{-12} N, the maximal stress component was 0.05 GPa, and the maximum displacement was 1×10^{-4} nm.

We calculated 10 parameters: the average number of electrons in s, p, and d orbits (N_s , N_p , and N_d); bond overlap population (BOP); heat of solution (ΔH_{relax}); heat of solution of ideal lattice ($\Delta H_{\text{unrelax}}$); misfit strain (ε); Young's modulus (E); shear modulus (G); and bulk modulus (K). N_s , N_p , N_d , and BOP were calculated using Mulliken population analysis.²⁰⁾ In Ag-, Cu-, and Al-based alloys, the values of BOP between the solute atom and solvent atom at the first nearest neighbor were used. In Mg-based alloys, the average value between the solute atoms and the solvent element at the first and second nearest neighbor was used. ΔH_{relax} and $\Delta H_{\text{unrelax}}$ were calculated using the following equations:

$$\Delta H = E_{\text{MX}} - (31/32 E_{\text{M}} + E_{\text{X}}) \quad (1)$$

$$\Delta H = E_{\text{MgX}} - (35/36 E_{\text{Mg}} + E_{\text{X}}) \quad (2)$$

where E_{M} is the energy of a solvent atom such as Ag, Cu, and Al. E_{Mg} is the energy of solvent atom Mg. E_{X} is the energy of a solute atom. When ΔH_{relax} was calculated, the energy after full structural relaxation was used. For $\Delta H_{\text{unrelax}}$, the energy before the structural relaxation was used. The misfit strain ε was calculated using the following equation:²¹⁾

$$\varepsilon = (d_{\text{MX}} - d_{\text{M}})/d_{\text{M}} \quad (3)$$

d_{M} is the interatomic distance in pure M such as, Ag, Cu, Al, and Mg, and d_{MX} is the first nearest neighbor distance between solvent atom M and solute atom X in M-X solid solutions. In the case of Mg-based alloys, the average values of the first and second nearest neighbor distances between solvent and solute atoms were used. Young's modulus (E),

shear modulus (G), and bulk modulus (K) were obtained by Hill's approximation.²²⁾

We used Neural Network Console, which is a deep learning tool for training, evaluating, and designing the artificial neural network. An optimization algorithm called adaptive moment estimation (ADAM)²³⁾ was used to calculate the gradient needed in calculation of weights associated with each connection between neurons.

The network structure was explored and optimized from the number of layers and various activation functions using the Bayesian optimization based on Gaussian processes.²⁴⁾ The network was composed of four or five layers including one input layer, two or three hidden layers, and one output layer. Huber loss function²⁵⁾ was used as the loss function. The Huber loss function could stabilize the training set better than the squared error function. We used the following activation functions. The Relu function²⁶⁾ is widely used to perform quick calculations. The Elu function²⁷⁾ is an improved form of the Relu function. Abs is an absolute function that transfers the absolute value to the next layer. The Tanh function²⁸⁾ is symmetrical around 0. The Swish function²⁹⁾ looks like a sigmoid function. The Selu function³⁰⁾ can undergo self-normalization. Batch normalization³¹⁾ is a process of normalizing within a mini-batch and can help in quick learning of the network. Dropout³²⁾ can prevent overfitting.

3. Results and Discussion

3.1 Data collection

In this study, we constructed two solid solubility limit prediction systems. One used Hume-Rothery parameters similar to the study of Zhang *et al.*,⁷⁾ while the other used results from the first-principles calculations. In the artificial neural network with the Hume-Rothery parameters, four explanatory variables were used: atomic radius difference between the solute atom and solvent atom, Pauling's electronegativity,³³⁾ which is the difference between the solute atom and the solvent atom, the valence of the solute element, and the valence of the solvent element. The number of valence electrons was obtained from the group number in the periodic table used by the Chemical Abstract Service (CAS).

We also used 11 explanatory variables: one was experimental data and the other ten were the results of the first-principles calculations. The experimental data were melting point of solute atoms T_{X} .³⁴⁾ The results of the first-principles calculations were N_s , N_p , N_d , BOP, ΔH_{relax} , $\Delta H_{\text{unrelax}}$, ε , E , G , and K in M-X solid solution alloys. Based on thermodynamic knowledge, the solid solubility limit depends on an excess Gibbs free energy of the solid solution, which can be subdivided into an enthalpy of mixing and entropy contribution. We expected that the results of the first-principles calculations in M-X solid solution alloys would be related to the excess Gibbs free energy of the solid solution, but the experimental melting point of solute atoms was especially expected to be related to the entropy contribution.

In any artificial neural network, we used the common logarithmic value of the solid solubility limit, c , collected from the equilibrium diagram^{35,36)} as the explained variable

Table 1 Dataset of Ag-based alloys.

Elements	$T_x^{34}) / K$	$\varepsilon / \%$	N_s / 1/atom	N_p / 1/atom	N_d / 1/atom	ΔH_{relax} / eV/atom	$\Delta H_{unlerax}$ / eV/atom	BOP	K / GPa	G / GPa	E / GPa	$\log c^{35, 36})$ / log(at.%)
Li	453.7	-0.785	1.72	0	0	-0.614	-0.899	-0.1	95.90	27.61	75.58	1.786
Be	1562	-2.533	2.47	1.48	0	1.016	0.877	-0.34	100.33	33.81	91.19	0.538
Mg	922	0.036	0.55	0.49	0	-0.698	-0.992	0.05	96.56	34.61	92.74	1.469
Al	933	-1.041	1.08	1.83	0	-0.271	-0.538	0.28	99.14	28.49	77.99	1.255
Sc	1814	0.385	0.38	0.66	1.51	-0.691	-0.974	0.18	96.11	30.18	81.96	0.995
Ti	1946	-0.799	0.45	0.78	2.56	0.619	0.343	0.22	98.10	33.60	90.46	0.766
Mn	1520	-1.956	0.34	0.49	5.81	2.571	2.390	0.18	102.80	33.79	91.36	1.670
Co	1770	-2.245	0.41	0.54	7.84	1.668	1.507	0.19	103.30	31.37	85.47	-0.114
Ni	1728	-2.141	0.43	0.6	8.84	0.770	0.589	0.18	101.52	28.34	77.78	0.107
Cu	1358	-1.615	0.64	0.85	9.79	0.333	0.100	0.2	100.66	30.66	83.50	1.154
Zn	692.7	-1.029	-1.08	1.37	9.97	-0.110	-0.382	-0.07	98.89	33.49	90.28	1.504
Ga	302.9	-0.487	-1.19	1.83	0	-0.103	-0.393	-0.21	97.93	29.48	80.37	1.257
Ge	1211	0.000	-1.23	2.44	0	0.338	0.043	-0.43	97.74	27.51	75.46	0.975
As	1089	0.377	-0.41	3.23	0	0.382	0.092	-0.33	97.95	31.21	84.63	0.916
Y	1801	2.410	0.35	0.21	1.68	-0.714	-0.819	0.12	96.54	29.75	80.93	0.132
Zr	2128	0.650	0.48	0.56	2.63	0.258	-0.013	0.22	99.69	33.44	90.23	-1.024
Mo	2895	-0.934	0.53	0.53	5.11	2.821	2.553	0.24	104.58	35.29	95.15	-0.835
Cd	594.3	0.776	-0.68	1.25	9.94	-0.198	-0.474	0.02	97.40	34.17	91.77	1.574
In	429.8	1.136	1.01	1.81	-0.02	-0.161	-0.415	0.2	96.99	28.28	77.33	1.305
Sn	505.1	1.482	0.69	2.41	0	-0.143	-0.366	0.12	97.15	26.54	72.98	1.061
Sb	903.9	1.969	0.4	3.11	0	0.113	-0.063	-0.01	96.80	29.15	79.48	0.841
La	1194	5.192	-2.01	-0.98	2.02	0.137	0.666	-0.46	91.88	24.84	68.35	-0.021
Pt	2045	-0.756	0.71	0.46	9.02	-0.115	-0.395	0.21	102.42	29.29	80.22	1.626
Hg	234.3	1.291	-1.03	1.24	9.88	0.112	-0.129	-0.12	97.26	33.19	89.40	1.577
Tl	577	2.119	-1.11	1.77	0.03	0.422	0.259	-0.25	95.13	27.85	76.12	0.457
Pb	600.7	2.951	-1.08	2.32	0.04	0.459	0.418	-0.43	94.01	25.47	70.08	-0.162
Bi	544.6	2.947	-0.63	2.98	0	0.107	0.080	-0.39	95.12	28.96	78.88	0.024

$\log c$. A total of 104 data including those for 27 Ag-based, 26 Cu-based, 30 Al-based, and 21 Mg-based alloys were collected. Datasets of the results of the first-principles calculations and the explained variables are shown in Tables 1, 2, 3, and 4. All explanatory variables and explained variables were normalized before training to make them fall within the range of -1 to 1 .

3.2 Machine learning

Figure 2 shows the optimized network that used the Hume-Rothery parameter. Figure 3 shows the prediction of the solubility using Hume-Rothery parameters in comparison with the experimental data. We used the hold-out method for the network: 80% of all data were training data and 20% of all data were validation data. The coefficient of determination R^2 for the correlation between predicted and experimental values was 0.715. The coefficient of determination of the study of Zhang *et al.* was 0.590.⁷⁾ The coefficient of determination in this work was better than that by Zhang *et al.* The prediction accuracy was improved because of the increase in the number of data of Al-based and Mg-based alloys.

Figure 4 shows the optimized network that used the results of the first-principles calculations. Figure 5 shows the prediction of solubility using the results of the first-principles calculations in comparison with experimental data. Among all data, 80% were training data and 20% were validation data, as shown in Fig. 5. The coefficient of determination, R^2 , for the correlation between predicted and experimental values was 0.900. Thus, the prediction system of solid solubility limit was more accurate than that constructed by Zhang *et al.*⁷⁾

Stepwise regression is a method of fitting regression models in which the explanatory variables are chosen by an automatic procedure. We used the stepwise method to optimize the number of explanatory variables of the network which used the results of the first-principles calculation. Figure 6 shows the principle diagram of the stepwise method used in this work. First, the explanatory variables are reduced by 1 from 11 explanatory variables, and the combination that maximizes R^2 is searched. Furthermore, the explanatory variables are reduced by one, and the optimal combination is searched. The hold-out method was used while selecting the

Table 2 Dataset of Cu-based alloys.

Elements	$T_x^{(34)}$ / K	ε / %	N_s / 1/atom	N_p / 1/atom	N_d / 1/atom	ΔH_{relax} / eV/atom	$\Delta H_{\text{unrelax}}$ / eV/atom	BOP	K / GPa	G / GPa	E / GPa	$\log c^{(3,5,3,6)}$ / log(at.%)
Li	453.7	0.346	1.14	0	0	-0.198	-0.215	-0.37	135.36	53.12	140.94	1.299
Be	1562	-1.715	2.1	1.56	0	0.108	0.165	-0.92	140.08	53.84	143.19	1.214
Mg	922	2.371	0.27	-0.18	0	-0.100	0.024	-0.13	142.79	54.51	145.06	0.840
Al	933	0.687	0.83	1.78	0	-0.693	-0.698	0.29	136.57	53.57	142.12	1.197
Si	1687	-0.237	1.28	2.53	0	-0.187	-0.204	0.3	140.70	50.49	135.29	1.039
P	863	0.059	1.64	3.29	0	0.142	0.123	0.24	127.01	46.85	125.17	0.509
Sc	1814	3.396	-0.01	0.08	1.61	-0.290	0.022	0	125.55	46.69	124.61	-0.062
Ti	1946	1.466	0.16	0.37	2.59	0.121	0.172	0.12	126.01	50.91	134.60	0.868
V	2202	0.429	0.22	0.36	3.68	1.087	1.077	0.15	143.05	51.60	138.19	-0.709
Cr	2133	1.323	0.28	0.42	4.86	0.594	0.623	0.13	135.88	50.37	134.48	-0.167
Fe	1811	-0.954	0.26	0.36	6.81	1.290	1.298	0.19	147.07	52.86	141.63	0.428
Co	1770	-0.725	0.31	0.34	7.8	0.660	0.657	0.19	139.20	47.48	127.90	0.861
Zn	692.7	0.702	-1.36	1.4	9.94	-0.268	-0.273	-0.04	137.72	54.34	144.08	1.504
Ga	302.9	1.073	-1.69	1.93	-0.01	-0.341	-0.328	-0.18	137.49	51.59	137.56	1.203
Ge	1211	1.159	-1.96	2.53	0	-0.053	-0.034	-0.42	139.83	49.12	131.91	1.072
As	1089	1.790	-0.6	3.21	0	0.211	0.278	-0.21	137.50	48.34	129.81	0.876
Nb	2750	2.474	0.27	-0.21	3.99	1.135	1.332	0.14	147.16	49.80	134.25	-0.064
Ag	1235	2.837	0.53	0.4	9.69	0.434	0.657	0.18	143.92	55.93	148.54	0.688
Cd	594.3	3.902	-0.74	1.14	9.9	0.544	0.975	0.06	123.37	54.60	142.74	0.300
In	429.8	4.219	0.91	1.77	-0.04	0.471	0.985	0.21	148.83	50.60	136.36	0.978
Sn	505.1	4.380	0.68	2.41	0	0.303	0.888	0.18	124.59	48.17	128.02	0.892
Sb	903.9	4.519	0.37	3.03	0	0.549	1.112	0.07	113.61	45.04	119.36	0.746
Hf	2504	4.471	0.4	2.19	2.69	0.537	1.229	0.53	134.46	48.42	129.68	-0.404
Ir	2720	1.341	0.45	0.23	7.84	0.120	0.170	0.25	152.90	48.29	131.06	0.909
Hg	234.3	4.469	-1.45	1.19	9.81	0.834	1.442	-0.12	135.01	53.41	141.56	0.730
Tl	577	5.138	-1.74	1.82	10.04	1.330	2.078	-0.32	136.73	47.53	127.79	-0.554

Fig. 2 Architecture of the optimized artificial neural network using Hume-Rothery parameters.

explanatory variables by the stepwise method: 80% of all data were training data and 20% of all data were validation data. Figure 7 shows the changes in the coefficient of determination for only validation data with the decreasing number of explanatory variables. Table 5 shows the

Fig. 3 Prediction of solubility using Hume-Rothery parameters in comparison with experimental data.

optimized combination of each number of explanatory variables. The best combination of explanatory variables includes eight explanatory variables which were the average number of electrons in the s and d orbits (N_s and N_d), heat of

Table 3 Dataset of Al-based alloys.

Elements	$T_x^{(34)}$ / K	ε / %	N_s / 1/atom	N_p / 1/atom	N_d / 1/atom	ΔH_{relax} / eV/atom	$\Delta H_{\text{unrelax}}$ / eV/atom	BOP	K / GPa	G / GPa	E / GPa	$\log c^{(35, 36)}$ / log(at.%)
Li	453.7	-0.466	1.82	0	0	-0.322	-0.316	-0.09	72.10	10.26	29.40	1.149
Be	1562	-2.306	2.53	1.57	0	0.789	0.910	-0.55	75.50	17.47	48.66	-0.535
Na	371	2.234	1.92	0.09	0	0.671	0.788	-0.14	70.35	4.23	12.45	-2.121
Mg	922	1.345	0.57	0.6	0	0.123	0.165	-0.04	71.66	10.66	30.46	1.267
Si	1687	-0.706	1.58	2.59	0	0.430	0.441	0.11	74.36	38.17	97.78	0.187
Ca	1113	4.344	0.02	-0.02	1.21	0.128	0.537	0.07	68.16	17.13	47.43	0.145
Sc	1814	1.398	0.21	0.54	1.99	-1.085	-1.031	0.17	74.04	19.25	53.15	-0.673
Ti	1946	-0.896	0.29	0.64	3.01	-0.955	-0.930	0.22	76.02	22.63	61.76	-0.846
V	2202	-2.507	0.32	0.59	4.12	-0.364	-0.189	0.22	79.03	27.59	74.13	-0.497
Cr	2133	-1.836	0.42	0.78	5.17	-0.320	-0.243	0.17	75.80	27.04	72.50	-0.432
Mn	1520	-4.175	0.26	0.4	6.29	-0.445	-0.381	0.17	79.42	32.74	86.34	-0.206
Fe	1811	-4.432	0.32	0.62	7.17	-0.398	0.142	0.2	78.81	24.37	66.29	-1.522
Ni	1728	-2.960	0.5	0.76	8.98	-0.814	-0.586	0.19	80.10	33.17	87.45	-1.593
Cu	1358	-1.639	0.7	1.03	9.76	-0.077	-0.013	0.14	75.08	14.88	41.87	0.391
Zn	692.7	-0.231	0.11	1.45	9.96	0.174	0.175	0.11	73.55	10.51	30.09	1.827
Ga	302.9	0.527	0.36	1.94	-0.01	0.118	0.125	0.09	72.54	13.23	37.42	0.943
Ge	1211	0.599	0.51	2.57	0	0.327	0.335	-0.06	72.73	31.91	83.51	0.297
Se	494	1.844	0.62	4.29	0	0.484	0.559	-0.69	71.46	28.01	74.32	-2.512
Y	1801	4.257	0.17	0.26	2.34	-0.639	-0.182	0.11	71.21	18.50	51.07	-1.287
Zr	2128	1.449	0.27	0.42	3.28	-1.038	-0.971	0.2	75.76	16.35	45.75	-0.767
Nb	2750	-0.701	0.32	0.41	4.57	-0.716	-0.693	0.2	79.99	28.96	77.52	-1.175
Ag	1235	0.442	0.67	0.94	9.74	0.052	0.058	0.1	74.92	10.49	30.07	1.531
Cd	594.3	2.359	0.19	1.39	9.93	0.645	0.779	0.11	71.78	1.44	4.29	-1.051
In	429.8	3.256	1.12	1.9	-0.02	0.776	1.037	0.21	70.64	8.64	24.91	-1.324
Sn	505.1	3.726	1.01	2.49	0	0.745	1.081	0.15	70.63	24.12	64.96	-3.521
La	1194	7.374	-0.76	-0.29	2.76	0.315	1.666	-0.31	67.96	21.79	59.07	-2.020
Hf	2504	1.099	0.36	0.16	3.27	-0.889	-0.846	0.19	76.56	17.84	49.65	-0.732
Ta	3293	-0.631	0.41	0.54	4.35	-0.678	-0.657	0.27	79.49	28.33	75.97	-1.399
W	3695	-2.399	0.62	0.69	5.29	-0.039	0.157	0.31	80.85	29.41	78.70	-1.617
Au	1338	0.367	1.02	1.08	9.59	-0.623	-0.618	0.11	75.82	16.54	46.27	-1.303

solution (ΔH_{relax}), heat of solution of ideal lattice ($\Delta H_{\text{unrelax}}$), Young's modulus (E), shear modulus (G), bulk modulus (K), and melting point of solute atoms (T_x).

Generally, machine learning faces a problem called overfitting.³²⁾ Therefore, we tested 10-fold cross validation to evaluate the generalization performance of the constructed network with eight explanatory variables. Figure 8 shows the principles diagram of 10-fold cross validation. The procedure of 10-fold cross validation is as follows. All data were divided into 10 groups: one was the validation data group and the others were training data groups. The network was trained using the training data, and the validation data were estimated. This process was repeated 10 times. Finally, R^2 was evaluated between experimental and all predicted values. Figure 9 shows the result of 10-fold cross validation. R^2 of the network that used eight explanatory variables was 0.6993 after 10-fold cross validation.

After 10-fold cross validation, we calculated the mean absolute error (MAE) of the network that used eight

explanatory variables of first-principles calculation results. Table 6 shows the MAE of each alloy obtained by the prediction network. For all alloys, the MAE of the network that used eight explanatory variables by the first-principles calculations was 0.45. Because the common logarithmic value was used as the explained variable, the solid solubility limit predicted from the network on average was $10^{-0.45}$ to $10^{0.45}$ times the true value. In other words, the solid solubility limit predicted from this network was on average 0.35 to 2.85 times the true value.

The accuracy of the solid solubility limit prediction system constructed in this study depended on the type of solvent elements as shown in Table 6. The MAEs of Ag-, Cu-, and Mg-based alloys were better than the value of 0.45 for all alloys. However, the MAE of Al-based alloys was 0.63, clearly inferior to the MAE for all alloys. Only Mg-based alloys were hcp structure, but the differences in crystal structures did not affect the accuracy of the prediction. Ag, Cu, and Mg have one or two valence electrons, and are

Table 4 Dataset of Mg-based alloys.

Elements	$T_x^{34)}$ / K	ε / %	N_s / 1/atom	N_p / 1/atom	N_d / 1/atom	ΔH_{relax} / eV/atom	$\Delta H_{\text{unrelax}}$ / eV/atom	BOP	K / GPa	G / GPa	E / GPa	$\log c^{3,5, 3,6)}$ / log(at.%)
Li	453.7	-0.232	1.87	0	0	-0.243	-0.604	-0.025	37.57	22.75	56.79	1.236
Al	933	-0.836	1.25	1.98	0	0.094	-0.254	0.29	39.04	21.72	54.98	1.082
Ca	1113	3.383	0.13	-0.01	0.9	0.036	-0.146	0.08	37.07	17.33	44.98	-0.107
Sc	1814	0.574	0.34	0.54	1.75	-0.155	-0.509	0.275	38.87	21.51	54.47	1.181
Ti	1946	-1.449	0.4	0.59	2.8	0.630	0.327	0.305	38.86	23.83	59.35	-0.918
Mn	1520	-4.789	0.34	0.31	6.07	1.582	1.681	0.25	23.19	24.25	53.95	-0.006
Zn	692.7	-1.437	-1.1	1.41	9.97	0.007	-0.316	-0.05	38.70	21.56	54.55	0.371
Ga	302.9	-0.780	-0.97	1.96	0	-0.238	-0.590	-0.2	38.70	22.37	56.28	0.496
Sr	1042	5.217	0.03	-0.03	1.36	0.531	0.361	0.095	36.16	16.12	42.11	-1.515
Y	1801	2.605	0.29	0.13	2.06	-0.161	-0.396	0.23	38.25	20.48	52.13	0.532
Zr	2128	0.034	0.38	0.29	3.05	0.243	-0.106	0.295	39.63	18.74	48.57	0.011
Ag	1235	-1.750	0.72	0.65	9.79	-0.458	-0.752	0.265	39.51	21.75	55.14	0.579
In	429.8	0.808	0.3	1.99	-0.02	-0.223	-0.586	0.15	38.66	21.13	53.63	1.277
Sn	505.1	1.359	-0.12	2.64	0	-0.252	-0.593	-0.06	38.92	22.57	56.75	0.482
Ba	1002	6.643	-0.66	0.13	1.78	1.000	1.400	-0.21	35.03	16.15	42.00	-2.711
La	1194	4.760	-2.12	-1.01	2.52	0.086	0.145	-0.415	36.50	11.70	31.71	-0.857
Au	1338	-2.462	1.06	0.78	9.65	-1.321	-1.542	0.28	40.12	22.30	56.45	-1.003
Hg	234.3	-0.303	-0.88	1.38	9.88	-0.513	-0.877	-0.05	38.84	21.32	54.07	0.203
Tl	577	1.431	-0.8	1.96	0.02	-0.034	-0.379	-0.16	38.20	21.57	54.45	1.181
Pb	600.7	2.377	-0.79	2.61	0.03	-0.034	-0.313	-0.375	38.45	21.80	55.00	0.884
Bi	544.6	2.644	-0.64	3.3	0	-0.091	-0.339	-0.555	37.74	18.51	47.73	0.098

Fig. 4 Architecture of the optimized artificial neural network using results of first-principles calculations.

simple metals which are metals with free-electron-like valence band structure. However, Al has three valence electrons and is a neighbor of Si in the periodic table. The

Fig. 5 Prediction of solubility using first-principles data in comparison with experimental data.

several first-principles studies have shown that directional bonds with covalent characters can be locally formed for Al atoms compared to Cu and Mg atoms.^{37,38)} It can be said that Al has a character intervening between simple metals and covalent materials, and that Al can form bonds with both metallic and covalent characters according to its environ-

Fig. 6 Principle diagram of stepwise method.

Fig. 7 Change in coefficient of determination for validation data decreasing number of explanatory variables.

Table 5 Combination of explanatory variables with the largest coefficient of determination for validation data.

Number of explanatory variables	Explanatory variables	R^2
1	ΔH_{relax}	0.2707
2	$\Delta H_{\text{relax}}, K$	0.4590
3	$\Delta H_{\text{relax}}, G, K$	0.6659
4	$\Delta H_{\text{relax}}, E, G, K$	0.6898
5	$N_{\text{ss}}, \Delta H_{\text{relax}}, E, G, K$	0.7735
6	$N_{\text{ss}}, \Delta H_{\text{relax}}, E, G, K, T_{\text{x}}$	0.7551
7	$N_{\text{ss}}, N_{\text{d}}, \Delta H_{\text{relax}}, E, G, K, T_{\text{x}}$	0.7961
8	$N_{\text{ss}}, N_{\text{d}}, \Delta H_{\text{relax}}, \Delta H_{\text{unrelax}}, E, G, K, T_{\text{x}}$	0.8438
9	$N_{\text{ss}}, N_{\text{d}}, \Delta H_{\text{relax}}, \Delta H_{\text{unrelax}}, \epsilon, E, G, K, T_{\text{x}}$	0.8247
10	$N_{\text{ss}}, N_{\text{p}}, N_{\text{d}}, \Delta H_{\text{relax}}, \Delta H_{\text{unrelax}}, \epsilon, E, G, K, T_{\text{x}}$	0.7331
11	$N_{\text{ss}}, N_{\text{p}}, N_{\text{d}}, \text{BOP}, \Delta H_{\text{relax}}, \Delta H_{\text{unrelax}}, \epsilon, E, G, K, T_{\text{x}}$	0.6950

ment. This tendency of Al atoms bonding may be related to the difficulty of predicting the solid solubility limit in Al-based alloys.

The prediction by the network constructed in this work was more accurate than that predicted by Zhang *et al.*⁷⁾ Thus, the constructed prediction system is useful for designing and development of new alloys with unknown solid solubility limit.

Table 6 Prediction accuracy of $\log c$ after 10-fold cross validation.

	Ag	Cu	Al	Mg	Al1
MAE	0.39	0.34	0.63	0.43	0.45
$10^{-\text{MAE}}$	2.44	2.16	4.30	2.72	2.85
$10^{-\text{MAE}}$	0.79	0.46	0.23	0.37	0.35

Fig. 8 Principle diagram of 10-fold cross validation.

Fig. 9 Comparison between experimental data and the predicted data using 8 explanatory variables after 10-fold cross validation.

4. Conclusions

Using artificial neural network and first-principles calculation results, we constructed a solid solubility limit prediction system with high prediction accuracy. The obtained results are shown below:

- (1) In Ag-, Cu-, Al-, and Mg-based alloys, the determination coefficient R^2 was improved from 0.715 to 0.900 using the first-principles calculation results as explanatory variables compared with the case of using the Hume-Rothery parameters.
- (2) The best combination of explanatory variables was searched and eight explanatory variables were obtained: average number of electrons in s and d orbit (N_s and N_d), heat of solution (ΔH_{relax}), heat of solution of ideal lattice ($\Delta H_{\text{unrelax}}$), Young's modulus (E), shear modulus (G), bulk modulus (K), and melting point of solute atoms (T_x).
- (3) Ten-fold cross validation was performed to evaluate the generalization of the network. The solid solubility limit predicted from this network was on an average 0.35 to 2.85 times the true value.

Acknowledgments

This study was supported by the Light Metal Educational Foundation, Inc.

REFERENCES

- 1) K. Kimura, H. Kushima, K. Yagi and C. Tanaka: *Tetsu-to-Hagané* **81** (1995) 757–762.

- 2) E. Clouet, J.M. Sanchez and C. Sigli: *Phys. Rev. B* **65** (2002) 094105.
- 3) W. Hume-Rothery, G.W. Mabbott and K.M.C. Evans: *Philos. Trans. R. Soc. London, Ser. A* **233** (1934) 1–97.
- 4) W. Hume-Rothery, P.W. Reynolds and G.V. Raynor: *J. Inst. Met.* **66** (1940) 191–207.
- 5) U. Mizutani: *Materia Japan* **48** (2009) 119–125.
- 6) S. Haykin: *Neural Networks and Learning Machines*, 3rd ed., (Pearson, New York, 2008).
- 7) Y.M. Zhang, S. Yang and J.R.G. Evans: *Acta Mater.* **56** (2008) 1094–1105.
- 8) D. Minami, T. Uesugi, Y. Takigawa and K. Higashi: *Int. J. Mod. Phys. B* **33** (2019) 1950055.
- 9) T. Shiraiwa, Y. Miyazawa and M. Enoki: *Mater. Trans.* **60** (2019) 189–198.
- 10) S. Miyazaki, M. Kusano, D.S. Bulgarevich, S. Kishimoto, A. Yumoto and M. Watanabe: *Mater. Trans.* **60** (2019) 561–568.
- 11) S. Wu, J. Yang and Z. Liu: *Mater. Trans.* **61** (2020) 691–699.
- 12) M. Shirai and H. Yamada: *Mater. Trans.* **61** (2020) 176–180.
- 13) N. Hasegawa, S. Hasegawa, T. Kitaoka and H. Ohtsu: *Mater. Trans.* **60** (2019) 758–764.
- 14) M.D. Segall, P.J.D. Lindan, M.J. Probert, C.J. Pickard, P.J. Hasnip, S.J. Clark and M.C. Payne: *J. Phys. Condens. Matter* **14** (2002) 2717–2744.
- 15) P. Hohenberg and W. Kohn: *Phys. Rev.* **136** (1964) B864–B871.
- 16) W. Kohn and L.J. Sham: *Phys. Rev.* **140** (1965) A1133–A1138.
- 17) D. Vanderbilt: *Phys. Rev. B* **41** (1990) 7892–7895.
- 18) H.J. Monkhorst and J.D. Pack: *Phys. Rev. B* **13** (1976) 5188–5192.
- 19) T.H. Fischer and J. Almlof: *J. Phys. Chem.* **96** (1992) 9768–9774.
- 20) R.S. Mulliken: *J. Phys. Chem.* **56** (1952) 295–311.
- 21) T. Uesugi and K. Higashi: *Comput. Mater. Sci.* **67** (2013) 1–10.
- 22) R. Hill: *J. Mech. Phys. Solids* **11** (1963) 357–372.
- 23) D.P. Kingma and J. Ba: arXiv:1412.6980 (2014).
- 24) C. Rasmussen and C. Williams: *Gaussian Processes for Machine Learning*, (The MIT Press, Cambridge, 2006).
- 25) V. Katkovnik: *IEEE Trans. Signal Process.* **46** (1998) 3104–3109.
- 26) Y. LeCun, Y. Bengio and G. Hinton: *Nature* **521** (2015) 436–444.
- 27) D.A. Clevert, T. Unterthiner and S. Hochreiter: arXiv:1511.07289 (2015).
- 28) X. Glorot and Y. Bengio: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, ed. by T. Yee Whye and T. Mike, (PMLR, Proceedings of Machine Learning Research, 2010) Vol. 9, pp. 249–256.
- 29) P. Ramachandran, B. Zoph and Q.V. Le: arXiv:1710.05941 (2017).
- 30) G. Klambauer, T. Unterthiner, A. Mayr and S. Hochreiter: Proceedings of the 31st International Conference on Neural Information Processing Systems, (Curran Associates Inc., Long Beach, California, USA, 2017) pp. 972–981.
- 31) S. Ioffe and C. Szegedy: arXiv:1502.03167 (2015).
- 32) N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov: *J. Mach. Learn. Res.* **15** (2014) 1929–1958.
- 33) L. Pauling: *J. Am. Chem. Soc.* **69** (1947) 542–553.
- 34) C. Kittel and P. McEuen: *Introduction to Solid State Physics*, (J. Wiley, Hoboken, N.J., 2005).
- 35) B. Predel: *Phase Equilibria of Binary Alloys*, (Springer-Verlag, Berlin Heidelberg, 2003).
- 36) H. Baker: *ASM Handbook Volume 3, Alloy Phase Diagrams*, (ASM International, Ohio, 1992).
- 37) T. Uesugi, M. Kohyama and K. Higashi: *Phys. Rev. B* **68** (2003) 184103.
- 38) S. Ogata, J. Li and S. Yip: *Science* **298** (2002) 807–811.