Machine-Learning Clustering Technique Applied to Powder X-Ray Diffraction Patterns to Distinguish Compositions of ThMn$_{12}$-Type Alloys

Keishu Utimula,* Rutchapon Hunkao, Masao Yano, Hiroyuki Kimoto, Kenta Hongo, Shogo Kawaguchi, Sujin Suwanna, and Ryo Maezono

A clustering technique is applied using dynamic-time-wrapping (DTW) analysis to X-ray diffraction (XRD) spectrum patterns in order to identify the microscopic structures of substituents introduced into the main phase of magnetic alloys. The clustering technique is found to perform well, identifying the concentrations of the substituents with success rates of $\approx 90\%$. This level of performance is attributed to the capability of DTW processing to filter out irrelevant information such as the peak intensities (due to the uncontrollability of diffraction conditions in polycrystalline samples) and the uniform shift of peak positions (due to the thermal expansion of lattices). The established framework is not limited to the system treated in this work, but is widely applicable to systems the properties of which are to be tuned by atomic substitutions within a phase. The framework has a broader potential to predict properties such as magnetic moments, optical spectra etc. from observed XRD patterns, by predicting such properties evaluated from predicted microscopic local structure.

1. Introduction

The concept of materials informatics based on big data science has attracted recent interest in the context of discovering and exploring novel materials.$^{[1]}$ Achieving high efficiency to acquire data, namely experimental measurements and analysis of materials, is necessary to accelerate the cycle of the exploration of materials. X-ray diffraction (XRD) analysis is quite commonly used to determine crystal structure dominating material properties.$^{[2]}$ The analysis has been accelerated by the increase in X-ray intensities as well as in the measurement environment.$^{[3]}$ Typical efforts to achieve analysis efficiency include such studies applying machine learning techniques to a series of XRD data in a systematic observation (e.g., dependency on concentration, temperature etc.) to extract significant information.$^{[4]}$
While materials informatics approaches combined with XRD data have been recently used to distinguish different phases (i.e., inter-phase identifications),\cite{5–11} no attempt has been made to address intra-phase phases thus far. The present study aims to provide a framework for predicting the concentrations of atomic substituents introduced into the main phase of polycrystalline magnetic alloys.

The ThMn₁₂-type (Figure 1) crystal structured SmFe₁₁₋ₓTiₓ has been regarded as one of the candidates for the main phase in rare-earth permanent magnets.\cite{12} The origin of intrinsic properties emerging at high temperature as well as that of the phase stability has not yet been clarified. Introducing Ti and Zr to substitute for Fe and Sm is found to improve the magnetic properties and the phase stability, as described in details in the Sec. Samples and experiments. To clarify the mechanism of how the substitutions improve the properties, it is desirable to identify and quantify substituted sites, preferably with high throughput efficiency to accelerate materials tuning. In this work, we have developed a machine learning clustering technique to distinguish powder XRD patterns and obtain microscopic identification of the atomic substitutions.

Ab initio calculations are used to generate supervising references for the machine learning of XRD patterns: We prepared several possible model structures with substituents located on different sites over a range of substitution fractions. Geometrical optimizations for each model results in slightly different structures. We generated many XRD patterns calculated from the atom substitutions. The machine learning technique for XRD patterns developed here therefore has a broader range of applications that is not limited only to magnetic materials, but further to those materials with properties tuned by atomic substitution.

Table 1. The numbers of inequivalent configurations of Sm₁₋ₓZrₓFe₁₂₋ₓTiₓ to be considered. The numbers in brackets indicate the structures constructed from the 2 × 2 × 2 supercell (Sm/Zr), while the rest are constructed from the 2 × 2 × 1 supercell (Fe/Ti).

| y/x | 0.0  | 0.3  | 1.0  | 1.3  | 2.0  |
|-----|------|------|------|------|------|
| 0.000 | 1 | 13 | 22 (1) | 27 | 61 |
| 0.125 | — | — | (2) | — | — |
| 0.250 | — | — | (7) | — | — |
| 0.375 | — | — | (6) | — | — |
| 0.500 | — | — | (10) | — | — |

Table 2. Space groups of the initial structures for the substitution models (Sm/Zr) at each concentration. For y = 0.5, for instance, there are six equivalent configurations of substituted sites in total. The number given in parenthesis represents the number of degenerate configurations within each symmetry of the initial structures for further lattice relaxations. Therefore, there are 26 configurations in total for generating simulated XRD patterns.

| y (space group/number of configurations) | 0.000 | 0.125 | 0.250 | 0.375 | 0.500 |
|-----------------------------------------|------|------|------|------|------|
| Imm2 (1) | 13 | 0 | 0 | 0 | 0 |
| Imm2 (2) | 0 | 13 | 0 | 0 | 0 |
| Imm2 (4) | 0 | 0 | 13 | 0 | 0 |
| Imm2 (1) | 0 | 0 | 0 | 13 | 0 |
| Imm2 (2) | 0 | 0 | 0 | 0 | 13 |
| Imm2 (4) | 0 | 0 | 0 | 0 | 0 |
| Cmm2 (2) | 27 | 27 | 27 | 27 | 27 |
| Cm (2) | 61 | 61 | 61 | 61 | 61 |
| Cm (2) | 61 | 61 | 61 | 61 | 61 |
| P1 (1) | 61 | 61 | 61 | 61 | 61 |
| P1 (1) | 61 | 61 | 61 | 61 | 61 |

When it is applied to the experimental XRD patterns, in addition to distinguishing atomic substitutions. The machine learning technique for XRD patterns developed here therefore has a broader range of applications that is not limited only to magnetic materials, but further to those materials with properties tuned by atomic substitution.

2. Results

For our target system, [Sm₁₋ₓZrₓFe₁₂₋ₓTiₓ], we examined the range for x and y as shown in Table 1, which is experimentally accessible. For a given concentration, several possible configurations for substituents exist. They are sorted into identical subgroups in terms of the crystalline symmetries, as described in the Experimental Section. Tables 2 and 3 summarize the possible space groups of substituted alloy structures (used as initial structures for computations) for given concentrations of Sm/Zr and Fe/Ti, respectively.

After applying lattice relaxations to the initial structures achieved by ab initio geometrical optimizations, we can calculate XRD patterns for the lattices. The procedure for the above are given in detail in the Experimental Section. We could therefore generate simulated XRD patterns as above, that is, 26 patterns for the Sm/Zr substitution, to be used as data for the clustering by unsupervised learning. We examine whether clustering can sort the patterns again correctly based on concentration.

The resultant XRD patterns (simulated) coincide fairly well with the experimental pattern, as shown in Figure 2. We see that...
The patterns have almost the same overall shape, with slight variations in the inter-peak distances depending on the substituent concentrations. DTW is expected to perform well to capture such slight variations, because: The method is designed to be applied to such signals given along an axis (e.g., time dependent signal, \( y(t) \)) so that it can extract only the shape of the signal ignoring uniform shifts along the axis. The method scores the dissimilarity between signals, \( i \) and \( j \), in terms of the DTW-distance, DTW\((i,j)\).

A clustering framework is generally specified as a combination of methods, \( a \odot b \), where “\( a \)”scores the dissimilarity,” and “\( b \)”links elements to form clusters based on the given dissimilarity.” In the present work, we employed the framework, [normalized constrained-DTW (NC-DTW)] \( \odot \) [Ward linkage method], using the NC-DTW and Ward method implemented in the “Scipy package.” Descriptions of linkage and dissimilarity-measure methods used in this work can be found in the Scipy documents, except for the DTW dissimilarity measures which were calculated by the fastDTW\(^{[14]} \) package. The framework is found to achieve clustering to distinguish the concentration of Sm/Zr substitutions with sufficiently high accuracy, 96.2% (one failure among 26 XRD patterns), as shown in Figure 3.

As the terming of machine-learning, we note that learning part is not for DTW \((a)\) but for the clustering part \((b)\), which is a form of unsupervised learning. It is an important factor, in general, how many data are used for the leaning, which is 124 (26) for Fe/Ti (Sm/Zr) substitutions, all used as the training data for the clustering as the normal terminology in the machine-learning. The number is small from the standard viewpoint of the machine-learning, but it is a common problem for Materials Informatics where each data is too expensive (taken by expensive and laboring experimental observations) to get larger number of data for the learning. One of the impact of the present work would be the fact that the clustering works fairly well even with such a small number of data since the DTW \((a)\)—a distance measure used in the present clustering algorithm \((b)\)—plays a great role in the feature space as a mapping of two different descriptors belonging to different substitution groups. That is, the DTW \((a)\) was found to successfully capture similarity/dissimilarity between the patterns belonging to the same/different groups. An XRD pattern forms a 12 000-dim. numerical data, \( \{I(2\theta)\}_{\theta=1}^{2,000} \), \((20 = 0–120 \text{ deg.}, \delta\theta = 0.01 \text{ deg.})\), as described in the Experimental Section in detail.

### 3. Discussion

#### 3.1. Limitation of DTW-Dissimilarity

When the same method (Ward \( \odot \) DTW) as in the Sm/Zr case is applied to Fe/Ti, the success rate for the recognition is reduced to 33.1%. We identify the reason for the success rate for the Fe/Ti being worse than that of Sm/Zr from the dependence shown in Figure 4. Since XRD reflects lattice constants as its peak position, we can take the unit cell volume, \( v \), as a representative quantity to be captured by the clustering recognition in a situation where the cell symmetry is kept unchanged. DTW dissimilarity, DTW\((i,j)\), can then be regarded to scale approximately to the difference of \( v \). The recognition can therefore be regarded as a framework to perform an inverse inference from the “difference of \( v \)” to identify the “difference of \( x \)” on the dependence of \( v(x) \), as shown in Figure 4. For Sm/Zr, “trace back mapping” from \( v \) to \( x \) is one-to-one, while for Fe/Ti this is not the case due to the degeneracy of \( v \) (many different values of \( v \) share the same \( x \)). Under such degeneracy, it is impossible to provide correct inferences of “difference in \( x \)” from a given “difference in \( v \).” This difficulty leads to the reduced success rate of Fe/Ti clustering recognition.

The problem can be resolved by exploiting the advantage of ab initio methods that they can provide several other quantities in addition to the optimized lattice parameters. Even when DTW\((i,j) \sim |v(x) - v(x)|\) is not effective because of the degeneracy in \( v(x) \), other quantities such as the magnetization \( M(v) \) may be non-degenerate (as shown in Figure 5) and hence represent a means to solve the degeneracy. Using magnetization is especially practical because both experimental and simulated values are available. We also note that the dependence in Figure 5

### Table 3. Space group of SmFe\(_{12-x}\)Ti\(_x\) with inequivalent sites of Ti substitutions. SmFe\(_{12}\) (I4/mmm) is used as an initial structure. The number given in parenthesis represents the number of degenerate configurations within each symmetry of the initial structures for further lattice relaxations. Therefore there are 124 configurations in total for generating simulated XRD patterns.

| \( x \) (space group/number of configurations) | I4/mmm (1) | Cm (1) | C2/m (2) | Imm2 (2) | Fmmm (2) |
|------------------------------------------|-------------------------------|--------|---------|---------|---------|
| 0.0                                      | Cm (1)                        | 0.5    | Cm (2)  | 1.0     | Cm (2)  |
| 1.0                                      | Cm (1)                        | 1.5    | Cm (2)  | 2.0     | Cm (2)  |

### Figure 2. Comparison of simulated XRD patterns (top) of SmFe\(_{12}\)Ti\(_x\) and experimental XRD patterns (bottom) of Sm\(_{1.05}\)Fe\(_{10}\)Ti\(_{1.25}\). The inset numbers are the main-phase peak positions.
Figure 3. Clustering over the XRD peak patterns (26 in total) of $[\text{Sm}_{1-y} \text{Zr}_y \text{Fe}_{11} \text{Ti}]$, performed by DTW (dynamic-time-wrapping) scoring and the Ward linkage method. Setting the threshold to 1000 for the dissimilarity (horizontal broken line), the patterns are clustered into four groups, sharing almost the same number of substitutions by Zr. The red arrows at the bottom shows errors where a zero substitution is wrongly sorted into the group with one substitution etc.

Figure 4. Dependencies of the unit cell volume ($v$) on the concentration $x$ for Fe/Ti (blue) and Sm/Zr (red) substitutions. Multiple plots with the same color and the same $x$ represent the different symmetries as given in Tables 2 and 3. Rectangular enclosures on the blue dependence show the degeneracy, that is, different $x$ may result in the same $v$.

(left panel) is consistent with the experimental finding\cite{15} that the magnetization per unit volume increases as the Zr concentration increases. By using a weight, $W(i,j) = (M(v_i) - M(v_j))^2$, we can revise the dissimilarity as

$$\text{Dissimilarity}(i,j) = \text{DTW}(i,j) \times W(i,j),$$

so that it can prevent from the problem due to the degeneracy. We have confirmed that the success rate is actually improved from 33.1% to 99.19%, as shown in Figure 6, by using the magnetization as a weighting.

3.2. How to Treat Experimental XRD

As shown in Figure 2, simulated XRD patterns ($s$) reproduce the experimental ones ($e$) well. The consistency is sufficient that direct
comparison to evaluate the DTW distance, DTW($\epsilon, s_j$), is appropriate for the clustering (it is usual to apply some pre-processing for law data (“e” or “s”) to determine corrections (“$\tilde{e}$” or “$\tilde{s}$”) to evaluate DTW($\tilde{e}, \tilde{s}$) in order to compensate for the difference between the idealized simulation and reality). By preparing simulated XRDs, $\{\{s\}_{i,j}\}_{\epsilon}$, in advance, we can identify such a $s_j$ for a given $\epsilon$ which gives the smallest distance, $|\epsilon - s_j|$. The simulated $s_j$ is accompanied by several quantities, $\{q_{\alpha}\}$, such as the formation energy, the magnetization, and the local geometrical configuration of substituents, evaluated by the ab initio method. The $\{q_{\alpha}\}$ can be the theoretical predictions for the observed $\epsilon$, serving as a machine-learning framework for XRD patterns assisted by ab initio simulations.

Figure 7 shows that such a distance, $|\epsilon - s_j|$, works fairly well, for an example of $\epsilon$ at a composition $\text{Sm}_{1.85}\text{Zr}_{0.0}\text{Fe}_{10.75}\text{Ti}_{1.25}$. For the general composition, $\text{Sm}_x\text{Zr}_y\text{Fe}_z\text{Ti}_t$, we can define the composition similarity between $\epsilon$ and $\{s_j\}$ as,

$$D = \sum_{\alpha=1}^{4} (e^{\alpha}_\epsilon - e^{\alpha}_{s_j})^2 .$$

In Figure 7, we see that DTW($\epsilon, s_j$) (ordinate) correlates well with the ‘composition similarity’. The closest $s_j$ giving the shortest DTW($\epsilon, s_j$) (black filled circle in the figure) has the closest composition to $\text{Sm}_{1.0}\text{Zr}_{0.0}\text{Fe}_{11.0}\text{Ti}_{1.0}$ than the other $\{s_j\}$. The prediction accuracy is improved by the increased number of simulation data, $\{\{s\}_{i,j}\}$. A straightforward way to increase the amount of simulation data is to use a more dense grid on $x$ but this requires a larger supercell and hence more computational power. For the present grid resolution, providing the best performance possible, the experimental XRD with Zr% = 0.0, 10.4, and 31.8 are identified to be the closest to Zr% (simulated) = 0.0, 25, and 37.5, respectively.

In the case with the degeneracy (Figure 4 for Fe/Ti substitution), the DTW distance is not capable of performing the clustering for $\{s_j\}$, and hence quite unlikely to be capable of identifying the closest $s_j$ for a given $\epsilon$ based on the $|\epsilon - s_j|$. The strategy of weighting by $W(i,j)$ (the magnetization) introduced in the previous section will not work in this case because for $\epsilon$ (experimental XRD patterns), the magnetizations are not always available. A possible solution to distinguish $\epsilon$ would be to use a plucked set, $A \subset A = \{s_j\}$, as follows: Since $A$ is generated by simulations, each element is accompanied with the quantities like the magnetization, the formation energy etc. By referencing the formation energies, we can pluck the degenerate candidates (e.g., $P$ and $Q$ in Figure 8) by excluding those with higher energies ($P$ in Figure 8) to form the plucked subset $A$. The degeneracy in $A$ is now excluded, and hence it is used as a pool of references to be identified as the closest $s_j$ to a given $\epsilon$ based on the DTW distance, $|\epsilon - s_j|$. The identified $s_j$ is accompanied by the physical quantities evaluated by the simulations, and hence they can be used as estimates for the sample corresponding to the experimental XRD, $\epsilon$.
the NC-DTW does not exhibit the best performance, a compromise is made because the peak-shifting properties of NC-DTW are required to treat the experimental data. We note that the simulated XRDs are reflecting structures at zero temperature while the experimental ones are subject to the thermal effects of finite temperature. Thermal effects lead to the broadening of peaks due to thermal vibrations as well as peak shifts due to thermal expansion. Since we are interested in the change within a phase (not inter-phase changes), the shifts are expected to be almost uniform, not to modify the inter-peak distances significantly because expansion occurs almost evenly for every degree of freedom of the lattice. Such uniform shifts are not detected by DTW by design, and hence, the scoring is not affected by thermal effects. This feature results in robustness against thermal noise in the experimental data, enabling direct comparison with simulation data at 0 K to evaluate $|c - s|$.

Based on the above observations, we deliberately chose DTW even though it does not achieve the best performance for simulated data as shown in Tables 4 and 5. Evidence for this argument has also been shown in the preceding study\(^7\) in which NC-DTW exhibits the best performance, among various techniques, to sort the various phases of experimental data.

As we addressed earlier, the appropriateness of using this method lies in its characteristic of being insensitive to parallel shifts and therefore sensitive to sensing only differences in the distance between peaks. If other representative methods in supervised machine-learning, such as SVM (Support Vector Machine) or ANN (Artificial Neural Network), are used to distinguish only the inter-peak distances, there would be no reason to conclude their performance being negative except that they require much larger number of training data than those used here, $\sim 150$. We note that the DTW itself is not the machine-learning technique though the total framework here combined with clustering ones is so. In this context, DTW is regarded as an excellent mapping (distance measure) between descriptors in the feature space characterizing the XRD shape. This is the reason for our successful achievement here even with such a little number of samples as $\sim 150$. In this sense, it is likely to get successful results when the DTW is applied to other supervised methods. For example, the DTW can be implemented in a SVM kernel function, though larger amount of data may be required.

Several works are known in which DTW is applied to analyze XRD patterns.\(^{[7,16]}\) While these studies applied DTW to distinguish phases (i.e., inter-phase works), the present study applies it to intra-phase identification. In the former works, DTW is used to distinguish major differences of peak positions occurring when the phase changes.\(^{[7]}\) In this study, on the other hand, we demonstrated a new capability of DTW, namely, it can distinguish even the far smaller changes of inter-peak distances that occur within a target phase. Using this capability, we can explore a new framework that enables identification of the microscopic geometries of the substituents introduced in a target phase assisted by machine learning technique.

### 3.4. Further Discussions

It is required to point out possible factors leading discrepancies for the predictions by the present scheme when it is applied to other general cases. In the present case, we could enjoy the good
fortune of the availability of computational conditions for DFT, especially on the choices of XC and pseudo potentials, those well describe experiments to create the reference data. Such a fortune is, however, not expected generally. It is well known that the DFT usually overestimates or underestimates lattice constants, depending on the conditions. In such cases, the prediction of the inter-peak distance gets to be incorrect, so the reference XRD patterns generated by the DFT calculation cannot be used as it is. Although the topological order for the peak appearances are identical, the inter-peak distances gets incorrect when compared to experiments. However, such a remedy would be expected to work to some extent that scales the horizontal axis of 2Θ properly so that the peak positions can be adjusted to get coincident to experimental ones. A couple of experimental XRD data as references could make it possible for such an adjustment scheme by using, for example, the least-square optimization to determine the optimal rate for the scaling.

Another ingredient for the discrepancy is that the experimental data may be affected by several factors not taken into account in the simulations of ideal crystals, such as thermal expansions as we mentioned before. The effects of grain boundaries in poly-crystal samples would lead to the broadening of the peaks. Since the performance of DTW analysis relies on how the peaks are clearly recognizable, such a broadening would be taken as a serious problem leading inaccurate predictions of the framework. However, this might be not the case because of the property of DTW, that is, the insensitivity to the horizontal shifts of signals: As far as the peaks are recognizable (otherwise, XRD analysis does not work in the first place), we can make them sharper by making the axis be contracted. Though the inter-peak distances are also modified by the contraction, the temporary shortening of distances is uniform and hence does not matter when we measure the similarity between patterns, for all of which we could apply the common rate of contraction. The contraction also introduces the shift of peak positions, but it is not detected by DTW because of the insensitivity to the horizontal shift.

We should as well address here the limitation of the applicability of the present framework. XRD simulations using DFT packages are getting quite common even for the researchers who are not mainly working on simulations but experimental and industrial domains. However, there are at least two factors we can point out that may still be limiting the applicability for the broader range of researchers. The first one is the computational costs for handling the models of doped systems which are the target problem of the present topic. To describe the doped systems by the conventional DFT, supercell structures with N-times enlarged unit-cells than those of pristine systems are required, for which the cost gets to be ~ N^3 times expensive. Not only for the computational costs, one has to consider vast numbers of possible configurations of the doped sites even under a fixed concentration, which sometimes gets beyond the feasible capacity of available memories. Another factor is the difficulty in the choice of DFT computational conditions as we addressed in the previous paragraph. We have to note that the present problem corresponds to a quite rare case where the conventional GGA works fairly well to capture the structure. It is more in general for such systems including d and f elements that careful choices or adjustments are required for XC such as DFT+U etc. In such cases, generating XRD references using DFT itself requires specialized knowledge about XC.

4. Conclusion

We have developed a clustering framework that can be applied to XRD patterns of alloys to distinguish the concentrations of the substituents. We found that the clustering works quite well to identify the concentrations when applied to the patterns of magnetic alloys based on SmFe12. Supercell models for the substitutions are found to work well with ab initio lattice relaxations, reproducing XRD patterns sufficiently well to coincide with experiment. The implementation of the clustering with [DTW dissimilarity scoring]© [Ward linkage method] is found to achieve a success rate of around 90% for determining substituent concentrations. The main reason for failure of clustering is identified as degeneracy, namely the situation where different concentrations result in almost the same lattice constant. By imposing quantities predicted by ab initio methods into the weighting function used for the dissimilarity scoring, such degeneracies are avoided to prevent clustering failure. Sufficiently good coincidence between simulated and experimental XRD patterns enables the framework to be used to predict unknown concentrations of the substituent introduced in the main phase of alloys from their XRD patterns. The established framework is applicable not only to the system treated in this work, but widely to systems, the properties of which, are to be tuned by atomic substitutions within a phase. The framework has a larger potential to predict properties from observed XRD patterns, not only concentrations, by predicting properties including magnetic moments, optical spectra etc. evaluated from predicted microscopic local structure (positions of substitutions etc.).

Experimental Section

Samples and Experiments: The X-ray diffraction (XRD) measurements for the powdered Sm-Fe-Ti were performed at beamline BL22B2 of SPring-8 (Proposal Nos. 2016B1618 and 2017A1602). The diffraction pattern of CeO2 was used to determine the X-ray energy of 25 keV. The diffraction intensities were collected using a sample rotator system and a high-resolution one-dimensional semiconductor detector (multiple MYTHEN system) with a step size of 2Θ = 0.006°. The samples were powdered from strip-cast Alloys, and the powder was put into a quartz capillary and encapsulated with a negative pressure of Ar gas.

Computational Details: To determine the structures of the target alloys, SmFe12−xZr−xTi, constructed a tetragonal (4/mmm) crystal structure of SmFe12 was constructed using experimental lattice parameters, a = 0.856 nm and b = 0.480 nm (b = c), of SmFe12, Ti [18] as an initial setting for further optimization. For Zr-substitutions replacing Sm sites (ranging from 1–4 atoms), a 2 x 2 x 2 supercell, containing 104 atoms in the primitive of tetragonal (Immm)2 of SmFe12, Ti (Figure 1) was constructed. All possible configurations were considered to cover the randomness of experimental substitutions, and some configurations were ignored by considering their symmetry, determined using FINDSYM software. Finally, only 26 supercells (Table 2) that possessed different space groups, Wyckoff site occupancies, were considered.

For ab initio calculations, the spin-polarized density functional theory (DFT) implemented in the Vienna ab initio simulation package (VASP) was used. [20–22] For systems such as the target that include transition metals and rare earth elements, it is generally known that the predictions
are critically influenced by the choice of exchange-correlation (XC) potentials used in DFT.[23–26] For the present case, it has been found that DFT+U is essentially inevitable if f-orbitals are treated as the valence range.[27–31] It has also been found that the generalized gradient approximation (GGA) works well if the 4f is treated as the core range described by pseudo potentials.[32–36] The revised Perdew–Burke–Erzerhof (RPBE)[37] functional for the GGA-XC was therefore used upon confirmation that the RPBE improves the optimized lattice parameters by resulting in values closer to experimental values than the PBE.[38] The pseudopotentials based on the projected augmented wave (PAW)[39] method were used.

The s and p semi-core states are included in the valence states, except for Sm, resulting in 12, 16, and 12 valence states for Zr, Fe, and Ti, respectively. The structural relaxations were performed until the force on each ion was smaller than 0.01 eV Å⁻¹. A plane-wave cutoff energy of 400 eV and a 5 × 5 × 5 Monkhorst–Pack grid was used which was large enough to obtain the convergence energy. Lattice relaxations with the above choices applied to SmFe₁₁Ti was confirmed to result in the lowest total energy with Ti at the 8i site, which is consistent with experiment[12,40] and ab initio calculation[41] of RFe₁₁Ti-type magnetic compounds. The optimized lattice parameters, a and c, were 0.851 and 0.473 [nm] are in good agreement with experiment.[18] These comparisons confirm our model is quite working well. With Ti substitution at the 8i site, the 14/mmm space group breaks and becomes Imm2 as shown in Figure 1.

Validation of Simulated XRD Patterns: Simulated XRD patterns were validated by comparison to experimental XRD patterns. The X-ray diffraction (XRD) patterns of the optimized structures were theoretically calculated by the powder diffraction pattern utility of the software VESTA.[42] The X-ray wavelength of 0.496 Å used in the experiment was used. The isotropic atomic displacement parameter (B) was set to 1.00 Å. Normalized XRD patterns having 2θ from 1 to 120 degrees at 0.01 degree intervals were obtained.

The simulated XRD pattern of SmFe₁₁Ti disagrees very well with the experimental XRD pattern of Sm₁₀Fe₁₁₀Ti₂₅ (Figure 2); however, the main-phase peak position of 13.41 deg. in the experiment is significantly shifted to 13.48 deg. in the simulation. This is because peak shift occurs if the lattice expands or contracts, and we found that the optimized lattice parameters from DFT are underestimated, accounting for the difference. Underestimated lattice parameters introduce systematic shift of peak positions only, the XRD profiles remains unchanged. When the Zr concentration increases, the main-phase peak position of the simulated XRD patterns shifts to larger 2θ, in accordance with the experimental results.

Hierarchical Clustering Analysis: Hierarchical clustering analysis (HCA) was used to identify the simulated XRD patterns. All clustering analysis was carried out using the Scipy package.[43] The descriptions of linkage and dissimilarity-measure methods being used in this work can be found in the Scipy documents, except for the DTW dissimilarity measures which were calculated by the fastDTW[44] package.

The package provides a variety of other methods than the present choice, DTW & Ward, as shown in Tables 4 and 5. The tables compare the performance achieved by each method for the identifications of Sm/Zr and Fe/Ti, respectively. The performance is evaluated in terms of the adjusted rand index (ARI), which measures the similarity between the true labels and the predicted labels with maximum and minimum scores of 1 and −1, respectively. The ARI calculations were performed by using the Scikit-learn package.[43] In the tables, several dissimilarity measures, NC-DTW, Euclidean, Cityblock, Cosine, and Correlation, with various linkage methods, Single, Complete, Average, Weighted, Centroid, Median, and Ward, have been compared. The ideal score of 1 is attained by all methods, except NC-DTW, in Sm/Zr structures. While the best method in Fe/Ti structures are cosine or correlation dissimilarity measure with Ward linkage each attaining a score of 0.53. The NC-DTW method provides lower performance than other methods in both structures since NC-DTW omits the peak-shift information while the rest are peak-position based dissimilarity measures. With the NC-DTW dissimilarity measure, the Ward linkage method exhibits good performance, with ARIs of 0.91 and 0.01 for Sm/Zr and Fe/Ti structures, respectively. Therefore, in this work, the focus is on NC-DTW with Ward linkage method.

Acknowledgements

The computations in this work have been performed using the facilities of the Research Center for Advanced Computing Infrastructure at JAIST. K.R.M. is grateful for the financial support of MEXT-KAKENHI (17H05478 and 16KK0097), FLAGSHIP2020 (project nos. hp180206 and hp180175 at K-computer), Toyota Motor Corporation, I-O DATA Foundation, and the Air Force Office of Scientific Research (AFOSR-AOARD/FA2386-17-1-0049). K.H. is grateful for the financial support of FLAGSHIP2020 (project nos. hp180206 and hp180175 at K-computer), KAKENHI grant (17K17762), a Grant-in-Aid for Scientific Research on Innovative Areas (16H06439), PRESTO (JPMJPR16NA) and the “Materials research by Information Integration Initiative” (MII) project of the Support Program for Starting Up Innovation Hub from Japan Science and Technology Agency (JST). R.H. is grateful for the financial support of the Development and Promotion of Science and Technology Talents Project (DPST) for a scholarship to study at the Faculty of Science, Mahidol University, and a research internship at JAIST.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

dynamic time warping, machine learning, magnetic alloys

Received: March 3, 2020
Revised: April 30, 2020
Published online: June 5, 2020

[1] H. Ikekata, K. Hongo, T. Isomura, R. Maezono, R. Yoshida, J. Comput.-Aided Mol. Des. 2017, 31, 379.
[2] K. Hongo, S. Kurata, A. Jomphoak, M. Inada, K. Hayashi, R. Maezono, Inorg. Chem. 2018, 57, 5413.
[3] S. Kawaguchi, M. Takemoto, K. Osaka, E. Nishibori, C. Morigyoshi, Y. Kubota, Y. Kuroiwa, Sugimoto, Rev. Sci. Instrum. 2017, 88, 085111.
[4] W. B. Park, J. Chung, J. Jung, K. Sohn, S. P. Singh, M. Pyo, N. Shin, K.-S. Sohn, IUCrJ 2017, 4, 486.
[5] A. G. Kusne, T. Gao, A. Mehta, L. Ke, M. C. Nguyen, K.-M. Ho, V. Antropov, C.-Z. Wang, M. J. Kramer, C. Long, I. Takeuchi, Sci. Rep. 2014, 4, 6367.
[6] S. K. Suram, Y. Xue, J. Bai, R. Le Bras, B. Rappazzo, R. Bernstein, J. Bjorck, L. Zhou, R. B. van Dover, C. P. Gomes, M. Greigore, ACS Comb. Sci. 2017, 19, 37.
[7] Y. Iwasaki, A. G. Kusne, I. Takeuchi, npj Comput. Mater. 2017, 3, 4.
[8] S. Li, Z. Xiong, J. Hu, Mater. Sci. Technol. 2018, 34, 315.
[9] V. Stanin, V. V. Vesselinov, A. G. Kusne, G. Antoszewski, I. Takeuchi, B. S. Alexandrov, npj Comput. Mater. 2018, 4, 43.
[10] H. Xing, B. Zhao, Y. Wang, X. Zhang, Y. Ren, N. Yan, T. Gao, J. Li, L. Zhang, H. Wang, ACS Comb. Sci. 2018, 20, 127.
[11] C. Oses, E. Gossert, D. Hicks, F. Rose, M. J. Mehl, E. Perim, I. Takeuchi, S. Sanvito, M. Scheffler, Y. Lederer, O. Levy, C. Toher, S. Curtarolo, J. Chem. Inf. Model. 2018, 58, 2477.
[12] K. Kobayashi, S. Suzuki, T. Kuno, K. Urushibata, N. Sakuma, M. Yano, T. Shoji, A. Kato, A. Manabe, J. Magn. Magn. Mater. 2017, 426, 273.
[13] E. Jones, T. Oliphant, P. Peterson, SciPy: Open source scientific tools for Python, 2001, http://www.scipy.org/.
[14] S. Salvador, P. Chan, Intell. Data Anal. 2007, 11, 561.
[15] T. Kuno, S. Suzuki, K. Urushibata, K. Kobayashi, N. Sakuma, M. Yano, A. Kato, A. Manabe, AIP Adv. 2016, 6, 025221.
[16] L. A. Baumes, M. Moliner, N. Nicoloyannis, A. Corma, CrystEngComm 2008, 10, 1321.
[17] K. Utimula, K. Nakano, G. I. Prayogo, K. Hongo, R. Maezono, SHRYa Suite for High-throughput generation of models with atomic substitutions implemented by python, 2019. arXiv:1911.08071.
[18] T. Kuno, S. Suzuki, K. Urushibata, K. Kobayashi, N. Sakuma, M. Yano, A. Kato, A. Manabe, AIP Adv. 2016, 6, 025221.
[19] H. T. Stokes, D. M. Hatch, J. Appl. Crystallogr. 2005, 38, 237.
[20] G. Kresse, J. Hafner, Phys. Rev. B 1993, 47, 558.
[21] G. Kresse, J. Furthmüller, Comput. Mater. Sci. 1996, 6, 15.
[22] G. Kresse, J. Furthmüller, Phys. Rev. B 1996, 54, 11169.
[23] K. Hongo, R. Maezono, J. Chem. Theory Comput. 2017, 13, 5217.
[24] T. Ichibha, Z. Hou, K. Hongo, R. Maezono, Sci. Rep. 2017, 7, 2011.
[25] K. Hongo, M. A. Watson, T. Itakata, A. Aspuru-Guzik, R. Maezono, J. Chem. Theory Comput. 2015, 11, 907.
[26] K. Hongo, N. T. Cuong, R. Maezono, J. Chem. Theory Comput. 2013, 9, 1081.
[27] P. Larson, I. I. Mazin, D. A. Papaconstantopoulos, Phys. Rev. B 2003, 67, 214405.
[28] S. Yehia, S. H. Aly, A. E. Aly, Comput. Mater. Sci. 2008, 41, 482.
[29] X. Liu, Z. Altounian, Comput. Mater. Sci. 2011, 50, 841.
[30] H. Pang, L. Qiao, F. S. Li, Phys. Status Solidi B 2009, 246, 1345.
[31] W. Cheng, S. Zhao, X. Cheng, X. Miao, J. Supercond. Novel Magn. 2012, 25, 1947.
[32] Y. Harashima, K. Terakura, H. Kino, S. Ishibashi, T. Miyake, J. Appl. Phys. 2016, 120, 203904.
[33] P. Larson, I. I. Mazin, D. A. Papaconstantopoulos, Phys. Rev. B 2004, 69, 134408.
[34] M. Duerrschmabel, M. Yi, K. Uestuener, M. Liesegang, M. Katter, H. J. Kleebe, B. Xu, O. Gutfleisch, L. Molina-Luna, Nat. Commun. 2017, 8, 54.
[35] A. Ismail, J. Hooper, J. B. Giorgi, T. K. Woo, Phys. Chem. Chem. Phys. 2011, 13, 6116.
[36] B. Puchala, A. Van der Ven, Phys. Rev. B 2013, 88, 094108.
[37] B. Hammer, L. B. Hansen, J. K. Nørskov, Phys. Rev. B 1999, 59, 7413.
[38] J. P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 1996, 77, 3865.
[39] G. Kresse, D. Joubert, Phys. Rev. B 1999, 59, 1758.
[40] Y. chang Yang, X. dong Zhang, L. shu Kong, Q. Pan, S. lin Ge, j. lian Yang, Y. fan Ding, B. sheng Zhang, C. tang Ye, L. Jin, Solid State Commun. 1991, 78, 313.
[41] Y. Harashima, K. Terakura, H. Kino, S. Ishibashi, T. Miyake, Phys. Rev. B 2015, 92, 184426.
[42] K. Momma, F. Izumi, IUCr, J. Appl. Crystallogr. 2011, 44, 1272.
[43] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, J. Mach. Learn. Res. 2011, 12, 2825.