Materials Today Communications 29 (2021) 102960

Contents lists available at ScienceDirect

Materials Today Communications

journal homepage: www.elsevier.com/locate/mtcomm

Material informatics for uranium-bearing equiatomic disordered solid solution alloys

He Huang [a,b,*], Xin Wang [c,**], Jie Shi [a], Huogen Huang [c], Yawen Zhao [c], Haiyan Xu [c], Pengguo Zhang [c], Zhong Long [c], Bin Bai [c], Tao Fa [c], Ce Ma [a], Fangfang Li [c], Daqiao Meng [c,**], Xiaoqing Li [b,**], Stephan Schönecker [b], Levente Vitos [b,d,e]

[a] Science and Technology on Surface Physics and Chemistry Laboratory, Mianyang 621907, PR China
[b] Applied Materials Physics, Department of Materials Science and Engineering, Royal Institute of Technology, Stockholm SE-10044, Sweden
[c] Institute of Materials, China Academy of Engineering Physics, Mianyang 621900, PR China
[d] Department of Physics and Astronomy, Division of Materials Theory, Uppsala SE-75120, Sweden
[e] Institute for Solid State Physics and Optics, Wigner Research Center for Physics, Budapest H-1525, Hungary

ARTICLE INFO

Keywords:
Uranium alloys
Disordered solid solution phase
Machine-learning model

ABSTRACT

Near-equiatomic, multi-component alloys with disordered solid solution phase (DSSP) are associated with outstanding performance in phase stability, mechanical properties and irradiation resistance, and may provide a feasible solution for developing novel uranium-based alloys with better fuel capacity. In this work, we build a machine learning (ML) model of disordered solid solution alloys (DSSAs) based on about 6000 known multi-component alloys and several materials descriptors to efficiently predict the DSSAs formation ability. To fully optimize the ML model, we develop a multi-algorithm cross-verification approach in combination with the SHapley Additive exPlanations value (SHAP value). We find that the ΔS_C, Λ, Φ_s, γ and $1/\Omega$, corresponding to the former two Hume − Rothery ($H-R$) rules, are the most important materials descriptors affecting DSSAs formation ability. When the ML model is applied to the 375 uranium-bearing DSSAs, 190 of them are predicted to be the DSSAs never known before. 20 of these alloys were randomly synthesized and characterized. Our predictions are in-line with experiments with 3 inconsistent cases, suggesting that our strategy offers a fast and accurate way to predict novel multi-component alloys with high DSSAs formation ability. These findings shed considerable light on the mapping between the material descriptors and DSSAs formation ability.

1. Introduction

Uranium alloys are considered as the primary nuclear fuel material for research and future commercial reactors owing to a combination of attractive properties, e.g., high thermal conductivity and fission atomic density, easy fabrication, and good compatibility with fuel cladding [1,2]. However, anisotropic growth and irradiation-induced swelling occurring in these alloys are still unresolved issues [1,3]. Considerable efforts on U-based alloys have been devoted to finding optimal compositions that possess body-centered cubic (bcc) structure (γ-phase) and improved mechanical and thermodynamic stability under the extreme environment of reactors [1–8].

Traditionally, the total weight of alloying elements added to the uranium matrix (abundance of ^{235}U∼5 at%) is limited to less than 10 wt%. This means that uranium is the only principal element. Recently, a commercial U-50 wt%Zr alloy (abundance of ^{235}U∼20 at%) was developed, which exhibits higher radiation-induced swelling resistance than U-10 wt%Zr [9]. Swelling of the U-50 wt%Zr fuel is expected to be around one volume percent per atom percent burn, approximately 15 times less than that of U-10 wt%Zr (∼15 vol percent per one atom percent burn). Inspired by this, we bring forth the concept of disordered solid solution alloys (DSSAs) with multi-principal elements into the design of new uranium alloys. It is well known that DSSAs with four or more principal elements (not containing uranium), firstly proposed by Yeh *et al.* [10] and Cantor *et al.* [11], have drawn much attention due to their outstanding performance in phase stability [12,13], mechanical properties [14,15], and irradiation resistance [16,17]. As the performance of DSSAs satisfy the demands of fuel materials, the development

* Corresponding author at: Science and Technology on Surface Physics and Chemistry Laboratory, Mianyang 621907, PR China
** Corresponding authors.
E-mail addresses: huanghe0913@126.com (H. Huang), wangxin@alum.imr.ac.cn (X. Wang), mengdaqiao@caep.cn (D. Meng), xiaoqli@kth.se (X. Li).

https://doi.org/10.1016/j.mtcomm.2021.102960
Received 22 August 2021; Accepted 30 October 2021
Available online 5 November 2021

of uranium-bearing DSSAs may provide a feasible solution for improving fuel performance.

It is worth noting that a proper method for designing uranium-bearing DSSAs should consider a large number of element combinations and broad compositional space before sample fabrication. An operative approach is filtering out the multi-component uranium alloys with low DSSAs formation ability and retaining those with high DSSAs formation ability through machine leaning (ML). As a data-driven method, ML has been extensively and successively applied to predict phase formation ability [18–20] and other desired properties [21–23] in various alloy systems. During the prediction process, the correlation and importance of features related to phase stability or other targeted properties are easily captured [20–22], which is very useful in order to clarify ambiguous or controversial issues, and unravel crucial parameters affecting the performance of materials.

To date, several empirical rules based on fundamental atomic and thermodynamic properties have been proposed [24–27–30] to distinguish or predict multi-component alloys with or without concentrated DSSP (i.e. DSSP or N(not)DSSP). These rules were generally obtained from the distribution of a certain amount of phase data on a two-dimensional parameter space. The combination of parameters and their ranges underlying these rules are very different and strongly dependent on the selected datasets. For example, Guo *et al.* [26] selected a narrow range of ΔH_{mix} ([−5, 5] kJ/mol) for DSSAs formation in multi-component alloys. Yang and Zhang [25] and Zhang *et al.* [28] proposed the combinations (Ω, δr) and (ΔH_{mix}, δr), respectively, to determine DSSP and NDSSP. The suggested δr ranges for solid-solution formation in these works are [0, 6.6] and [1,6], respectively. Furthermore, δr, ΔS_c, and ΔH_{mix} were suggested to determine DSSP and NDSSP by Guo *et al.* [24] and Zhang *et al.* [27]. Although the same parameters were selected in these two works, the ranges for DSSAs formation are very different, i.e., $\delta r \leq 8.5$, $11 \leq \Delta S_c \leq 19.5$ J/(K · mol), $-22 \leq \Delta H_{mix} \leq 7kJ/mol$ from Guo *et al.* [24] and $\delta r < 4$, 13.38 J/(K · mol) $< \Delta S_c$, $-10 < \Delta H_{mix} < 5kJ/mol$ from Zhang *et al.* [27]. Recent results [29,30] from a two-dimensional parameter space analysis showed that the predictability of some parameter pairs for DSSP formation are quite similar, for instance, (δd, γ) and ($\sqrt{\delta H^0_{mix}}$, $\sqrt{\delta H_{mix}}$). This may indicate that certain parameter pairs are strongly correlated. However, a quantitative description of the importance of each parameter and possible correlations among these is still lacking. A signature of this deficiency is the failure of these empirical rules, typically obtained from about one hundred experimental data, to predict new DSSAs [31].

The purpose of this paper is to build ML models to extract materials informatics of known DSSAs hitherto reported, and develop a fast and accurate ML model to screen new uranium-bearing, equiatomic alloys never known before (see Table S1 of supplementary material), thereby narrowing down the phase space for experimental and/or theoretical investigations. To test accuracy and reliability of the interpretable ML models, 20 uranium alloys were synthesized and characterized.

The paper is organized as follows. Section 2 presents our machine learning technique for building the DSSP formation ability models and the experimental details of alloy fabrication and characterization; section 3 is for results and discussion, and a summary is presented in section 4.

2. Methodology

2.1. Property database and material descriptors

In this work, we compiled ∽ 1000 DSSAs, which contains metals, alloys, intermetallic compounds (IM), metallic glasses (MG) (see Table S2 in supplementary material). We also used the MG database (6836 entries) built by L. Ward [18]. Possibly inconsistent phase structure information on an alloy across different publications was considered here; two rules for choosing data were established: (1) the phase structure information were taken from XRD, (2) the structure information consistent with the majority of previous studies on that alloy was preferred. Finally, our database contains almost all the DSSAs hitherto reported, constituted by 47 different elements. After purifying ∽ 8000 entries of this database, 5979 non-replicate records are retained to form our property database for building the ML DSSAs model. We classify these multi-component alloys into two classes, namely DSSP and NDSSP.

To train a DSSAs model by ML, a set of physically motivated but easily obtained descriptors is required. So far, many empirical rules based on fundamental atomic and thermodynamic properties have been proposed to predict the DSSAs as aforementioned. In this work, we choose 22 initial parameters [24,25,32–35] as material descriptors. These 22 descriptors are related to composition (c_i) and basic physical properties ($_{pi}$) of the constituent elements, such as their atomic radius (r), melting temperature (T_m), entropy (S), enthalpy (H), cohesive energy (E_c), valence electron concentration (*vec*), and electronegativity (χ). Their physical meaning is briefly described in Table 1.

2.2. Machine learning model

The procedure for predicting the DSSAs formation ability of multi-component alloys by ML is illustrated in Fig. 1. First, we need to address the sample bias in this database (NDSSP: 93.18%, DSSP: 6.82%). We randomly sample 500 NDSSP alloys 10 times and combine these data with DSSP alloys to construct 10 subdatasets (NDSSP: 55.07%, DSSP: 44.93%). Then, to address the issue of the algorithm bias, we use the random forest (RF), logistic regression (LR), support vector machine (SVM), neural network and ada boost classifier (Ada) algorithms, implemented in the Artificial Learning and Knowledge Enhanced Materials Informatics Engineering (ALKEMIE) [36], to evaluate the subdatasets. The 10-fold cross-validations with a splitting data of 80% training set and 20% testing set are applied to assess the effectiveness of each model and avoid over-fitting. We find that these subdatasets show better performance than the original dataset, as shown in Fig.S1. All these subdatasets show a little lower recall values (0.82–0.94) than that of original dataset (0.96–0.98). However, the Jaccard value (JValue) [37], JValue = (True Positive)/(True Negative + True Positive + False Positive), of 10 subdatasets are all significantly higher (> 0.8) than that of original dataset (0.58–0.67). These results reveal that our strategy address the sample bias and algorithm bias, thereby significantly improving the prediction ability of these 5 models.

To screen out the most important descriptors, for each subdataset, we evaluate all the descriptors using the RF, extreme gradient goosting (Xgboost), extra trees classifier (ET) and gradient boosting classifier (GBC) model in combination with the SHapley Additive exPlanations value, nameyly SHAP value [38], which is implemented in PyCaret [39]. In contrast to the feature importance, the SHAP value take the descriptors contribution into consideration, which is one of the state-of-the-art algorithms to screen descriptors. The details explanation can be found in ref [38]. Then, we average these SHAP values for each material descriptor, as shown in Fig. 2. Obviously, the ΔS_c is the most important descriptors, followed by Λ, Φ_s, γ, Ω^{-1}, δ_r, *Dr*, $\Delta H^{ij\_max}_{mix}$, T_m and E_c.

Furthermore, to evaluate the effect of the material descriptor on the models, we use these 4 ML models above again with 3, 5 and 10 most important descriptors. These 4 supervised ML models learn these estimated descriptors, and phase type (y_i) of the 908 alloys in each subdataset. In terms of JValue,the prediction accuracy of 4 ML models with 3, 5 and 10 descriptors is comparable to their counterparts with original descriptors (within ± 5% errors), as shown in Fig.S2. For some subdatasets, the new models show even better performance than the original models. Finally, taking the time cost and model complexity into consideration, we choose 5 most important descriptors, namely ΔS_c, Λ, Φ_s, γ and Ω^{-1} through the whole paper.

Table 1
The initially selected 22 features [24,25,32–35] for identifying DSSP and NDSSP of multi-component alloys.

Number	Parameter	Formula	Description
1	δr	$100\times\sqrt{\sum_{i=1}^{n}c_i(1-\frac{r_i}{r})^2}$	Difference of single atomic radius
2	δ**d**	$\sqrt{\sum_{i=1}^{n}\sum_{j>1}^{n}c_ic_j(1-\frac{r_i+r_j}{2r})^2}$	Difference of pair atomic radius
3	Dr	$\sum_{i=1}^{n}\sum_{j>1}^{n}c_ic_j\lvert r_i-r_j\rvert$	Mismath of single atomic radius
4	γ	$\frac{(1-\sqrt{\frac{(r_{min}+r)^2-r^2}{(r_{min}+r)^2}})}{(1-\sqrt{\frac{(r_{max}+r)^2-r^2}{(r_{max}+r)^2}})}$	Mismath of atomic packing
5	T_m	$\sum_{i=1}^{n}c_iT_m^i$	Average melting temperature
6	ΔS_c	$-R\sum_{i=1}^{n}c_i\ln c_i$	Configuration entropy
7	Φ*	$\frac{\Delta S_c-S_H}{\lvert S_E\rvert}$	Dimensionless thermodynamic parameter
8	Ω^{-1}**	$\frac{\Delta H_{mix}}{T_m\Delta S_c}$	Ratio of Gibbs free energy
9	E_c	$\sum_{i=1}^{n}c_iE_{c,i}$	Average cohesive energy
10	ΔH_{mix}	$\sum_{i=1}^{n}\sum_{j>1}^{n}4c_ic_jH_{mix}^{ij}$	Mixing enthalpy
11	$\Delta H_{mix}^{ij\_max}$	$\max(\Delta H_{mix}^{ij})$	Largest ΔH_{mix}^{ij}
12	$\Delta H_{mix}^{ij\_min}$	$\min(\Delta H_{mix}^{ij})$	Smallest ΔH_{mix}^{ij}
13	$\sqrt{\delta H_{mix}}$	$\sqrt{\frac{\sqrt{\sum_{i=1}^{n}\sum_{j>i}^{n}c_ic_j(H_{mix}^{ij}-\Delta H_{mix})}^{-2}}{k_BT_m}}$	Dimensionless mixing enthalpy
14	$\sqrt{\delta H_{mix}^{0}}$	$\sqrt{\frac{\sqrt{\sum_{i=1}^{n}\sum_{j>i}^{n}c_ic_j(H_{mix}^{ij}-0)^2}}{k_BT_m}}$	Dimensionless mixing enthalpy
15	$\sqrt{\delta H_{mix}^{0+}}$	$\sqrt{\frac{\sqrt{\sum_{i=1}^{n}\sum_{j>i}^{n}c_ic_j(H_{mix}^{ij+}-0)^2}}{k_BT_m}}$	Dimensionless mixing enthalpy
16	$\sqrt{\delta H_{mix}^{0-}}$	$\sqrt{\frac{\sqrt{\sum_{i=1}^{n}\sum_{j>i}^{n}c_ic_j(H_{mix}^{ij-}-0)^2}}{k_BT_m}}$	Dimensionless mixing enthalpy
17	**vec**	$\sum_{i=1}^{n}c_ivec_i$	Average valence electron concentration
18	Δ**vec**	$\sqrt{\sum_{i=1}^{n}c_i(vec_i-vec)^2}$	Difference of valence electron concentration
19	D*vec*	$\sum_{i=1}^{n}\sum_{j>1}^{n}c_ic_j\lvert vec_i-vec_j\rvert$	Mismatch of valence electron concentration
20	$\Delta\chi$	$\sqrt{\sum_{i=1}^{n}c_i(\chi_i-\chi)^2}$	Difference of electronegativity
21	Dχ	$\sum_{i=1}^{n}\sum_{j>1}^{n}c_ic_j\lvert\chi_i-\chi_j\rvert$	Mismatch of electronegativity
22	Λ***	$\frac{\Delta S_c}{\delta r^2}$	Effective configuration entropy

* The expressions for SH and SE are given in Refs. [32,34,35].
** The reciprocal of Ω is defined in order to avoid divergence when $\Delta H_{mix}=0$ (occurs for some alloys).
*** Singh *et al.* [33] originally proposed Λ as a geometrical parameter. Here, we interpret it as effective configuration entropy.

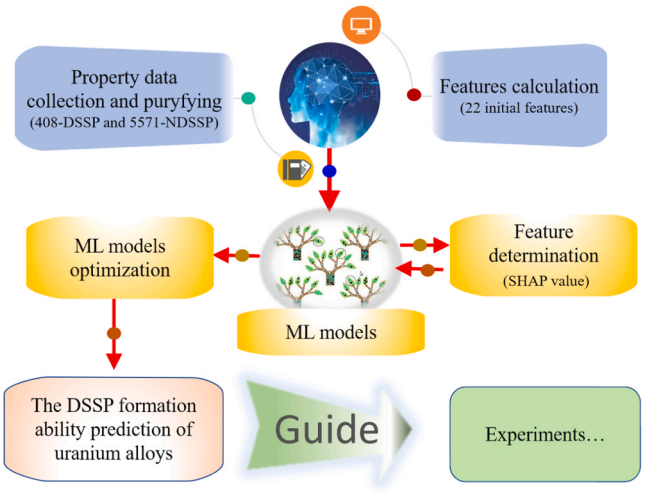

Fig. 1. Schematic for the ML prediction. First, the phase structure information for 5979 alloys and 22 initial material descriptors affecting DSSAs formation are collected from previous works. This database is divided into 10 subdatasets with 500 NDSSP and 408 DSSP alloys in each subdataset. Second, material descriptors are scored based on the analyses of SHAP values. Third, the DSSAs probability of uranium alloys is predicted by the ML models. Fourth, the uranium alloys with high DSSAs formation ability are selected as a guide to experiments.

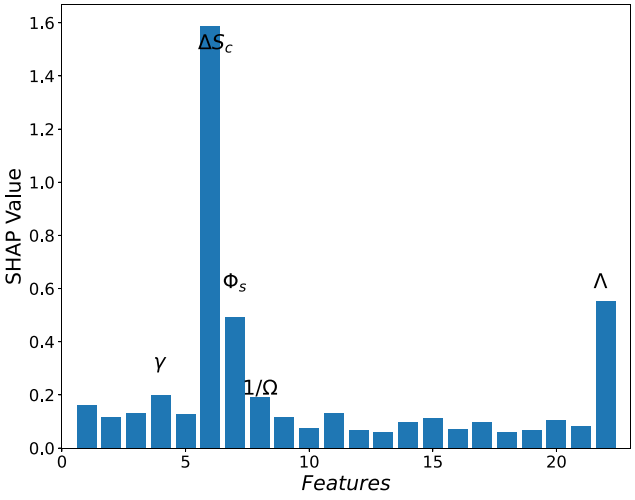

Fig. 2. Average SHAP values for descriptors. The descriptors enumerated from 1 to 22 correspond to that explained in the Table 1.

2.3. *Experimental processing*

Alloys were fabricated by vacuum arc-melting. The purity of uranium and all the other elements were 99.96 wt% and ≥ 99.8 wt%, respectively. Prior to melting, the furnace chamber was evacuated to 10^{-3} Pa and subsequently pressurized with Ar gas (purity, 99.999 vol%) to 0.5 Pa. All the alloys were remelted 4 times for homogeneity in the arc furnace equipped with a water-cooling copper mold.

Next, the cast ingots were cut into rectangular shape and polished for XRD measurement using SiC papers. The XRD patterns were acquired on an Empyrean diffractometer operated at 40 kV and 40 mA with Cu Kα radiation ($\lambda = 0.15406$ nm) and a diffracted beam monochromator. Data treatments including background and instrumental half width correction, $K\alpha_2$ stripping, and peak identification, were carried out using the Jade software package.

3. Results and discussions

Using these models trained by the four ML algorithms as discussed above, we predict DSSAs formation ability for 375 uranium-bearing equiatomic alloys. This set of alloys contains 45 ternary, 120 quaternary, and 210 quinary alloys, as shown in Table S1. We predict 190 novel uranium-bearing equiatomic alloys out of 375 candidates. For the experimental verification, 20 equiatomic alloys were randomly selected, fabricated by arc-melting and characterized by XRD. The phase information of these selected uranium alloys as-predicted and experimentally determined are listed in Table 2. The detailed XRD results for these alloys are shown in Fig.S3. The agreement is excellent with only three inconsistent cases.

It is very interesting to learn how the ML algorithms pick the candidates listed in Table S1 by investigating the physical descriptors. In Fig.S4 of the supplementary material, we plot the typical SHAP values of a subdataset for 5 most important descriptors. It clearly shows that ΔS_c, Λ and Φ_s present a positive effect on the model output, while the γ and $1/\Omega$ show a negative effect. The positive effect means the larger the value, the higher DSSAs formation ability. In Fig. 3, we plot the ΔS_c, Λ, Φ_s, $1/\Omega$ against γ of uranium-bearing DSSAs, together with those from our property database. For the predicted candidates, they show a similar ΔS_C as the DSSAs in the property database. Such a distribution is also observed on the Λ, Φ_s and $1/\Omega$. The ML indeed picks candidates, which have a higher ΔS_C, Λ and Φ_s, but a smaller γ and $1/\Omega$. In other words, the known high DSSAs formation ability in the property database are dominated by materials, which have a higher entropy contribution. Clearly, the ML picks candidates from the learned material information of the DSSAs in the property database.

Table 2
Phase information of 20 randomly selected uranium based equiatomic alloys obtained from prediction and experiment. The corresponding predicted DSSP probability and phase type of these alloys are also listed.

No.	Alloy	Predicted phase type	Experimental structure	Whether prediction agrees with experiment
1	UTiNbMoTa	DSSP	BCC1 +BCC2	Yes
2	UTiNbTa	DSSP	BCC1 +BCC2	Yes
3	UTiMoTa	DSSP	BCC1 +BCC2	Yes
4	UMoTa	DSSP	BCC1 +BCC2	Yes
5	UVNbMo	DSSP	BCC1 +BCC2	Yes
6	UTiVNbMo	DSSP	BCC1 +BCC2	Yes
7	UVNb	DSSP	BCC1 +BCC2	Yes
8	UTiCrNbMo	DSSP	BCC1 +BCC2 +IM	**No**
9	UNbMoHfTa	DSSP	BCC	Yes
10	UTiZrNbMo	DSSP	BCC	Yes
11	UNbMoHf	DSSP	BCC	Yes
12	UTiZrHf	DSSP	BCC+IM	**No**
13	UZrNbHf	DSSP	BCC	Yes
14	UTiCoNiMo	DSSP	Complex phases	**No**
15	UNbHf	DSSP	BCC	Yes
16	UTiCrZrNb	NDSSP	BCC+IM	Yes
17	UTiCrZrMo	NDSSP	Complex phases	Yes
18	UCrZrNbMo	NDSSP	Complex phases	Yes
19	UTiCoZrNb	NDSSP	Complex phases	Yes
20	UTiCoNiZr	NDSSP	Complex phases	Yes

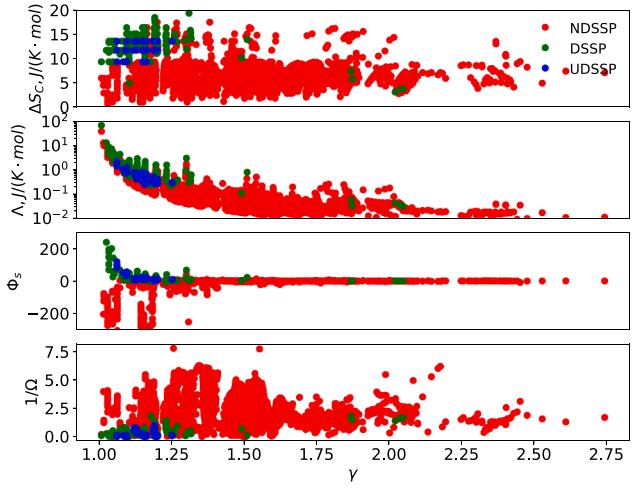

Fig. 3. ΔS_c, Λ, Φ_s, $1/\Omega$ and γ of uranium-bearing DSSAs, together with those from our property database.

It is interesting to observe in what way the informatics discovered from this property database can be related to the current understanding of solid solution formation mechanism. The design of DSSAs can date back to the Hume − Rothery ($H-R$) rules in physical metallurgy, which provide the key guidance to evaluating whether a binary solid solution can be formed [40]. These rules include: (i) the relative atomic size difference between the solute and solvent elements should be less than 15%; (ii) the formation of stable intermediate compounds should be restricted by carefully choosing the combination of metallic elements; and (iii) the electron concentration of the constituent elements should be tuned in favor of the formation of solid solutions.

As for the ΔS_c value, a positive effect feature, it helps to stabilize the solid solution phase over others, such as IM phase. For U-bearing DSSAs, it is only three values, corresponding to the values of ternary, quaternary and quinary alloys, respectively. In contrast, most of the NDSSP alloys show a lower ΔS_c value ($< 10\ J/(K \cdot mol)$). It has been pointed out that attaining high entropy in an alloy is not only compositionally dependent, as originally proposed, but also related to the atomic size/packing and formation enthalpy [41]. This explains why there are several atomic size difference factors, e.g. Λ and γ present in the 5 most important descriptors.

Considering the $\Lambda = \frac{\Delta S_c}{\delta r^2}$, we here interpret this parameter as effective configuration entropy (it was firstly proposed by Singh *et al.* [33] as a geometrical parameter). This interpretation intrinsically reveals the competition between misfit entropy and configuration entropy [42]. It is known that the effect of atomic size differences δr^2 on DSSAs formation is approximately proportional to the misfit entropy [42]. Thus, δr^2 could strengthen or weaken the effect of the configuration entropy ΔS_c. The new interpretation better distinguishes amorphous multi-principal element alloys from DSSAs characterized by an equal number of elements and the same set of chemical concentrations $\{c_i\}$ (but different chemical constituents), since solid solution and amorphous phases typically occur in different (though not disjoint) regions of atomic size difference [24]. Moreover, it is more obvious why DSSAs could be obtained by adjusting the composition or reducing the number of elements.

In terms of γ, a negative effect feature, most of the predicted U-bearing DSSAs show a lower value (1.0–1.25), in-line with that of DSSP alloys in our property database. From the viewpoint of physical foundation, the atomic packing parameter, γ, has a more distinct meaning, referring to the topological instability [40]. It presents the normalized packing density discrepancy between the smallest and largest atoms which predominantly determine the atomic packing instability. The smaller this value, the more stable is the alloy. Otherwise, IM phase may form.

Regarding Φ_s [41], a single dimensionless thermodynamic parameter, is related to the configurational entropy of mixing for an ideal gas (S_C), the excessive entropy of mixing (S_E), which is a function of atomic packing and atom size, and the complementary entropy (S_H) derived from the enthalpy of alloy. This parameter refers to the high entropy effect, namely the increase of Φ_s will maximize the entropy parameter, thus favoring the formation of DSSAs. This explains the positive effect of Φ_s as mentioned. In contrast, the negative Φ_s parameters indicate a diminishing entropy effect on the phase formation in these alloys. In our database, most of the MG present a negative value or near-zero value, which is consistent with previous findings [41].

Here it should be noted that these parameters are the necessary conditions for determining the solubility of multi-component alloys but not sufficient ones. These 5 important desriptors focus on the entropy contribution (S_C,S_E and S_H), which is related to the atomic size difference and atomic packing, corresponding to the H − R rules (i) and (ii). These results also evidence that the DSSAs formation ability is a synergy effect of these material descriptors.

Interestingly, we find that the predicted candidates can also form a dual-bcc phase, (e.g. UTiNbMoTa and UTiNbTa), besides single-bcc phase, e.g. UNbHf, as show in Table 2. We also interest in the formation mechanism of dual-phase and single-phase DSSAs. Previous investigations reveal that Φ_s [41], Λ [33] and *vec* [24] are three important descriptors, which can distinguish the single phase and multi-phase DSSAs. In Fig.S5, we plot the electronic concentration related parameters and Φ_s against Λ, we find that these three rules seem to have failed to distinguish these two classes when the data in database becomes larger. As for *vec*, there seems to be two groups for single-phase (SP) and multi-phase (MP) subdatasets with a "gap", as shown in top panel in Fig. S4. Regarding Φ_s, it cannot distinguish the SP ($\Phi_s > 20$) and MP DSSAs ($\Phi_s < 20$), as suggested in ref [40]. Interestingly, the Λ value of 0.24 $J/$()K·mol almost classify most of alloys to be the DSSAs, not the value 0.95, as suggested by Singh [33]. For clarity, we only plot the experimental results in Fig. 4. Surprisingly, we find that the Φ_s criterion, namely Φ_s=25, and Λ criterion, namely Λ=0.7 $J/(K \cdot mol)$, can distinguish the U-bearing dual-phase (UDP) from the U-bearing SP (USP) DSSAs with only one inconsistent case, as shown in Fig. 4. Here we note that this inconsistent case is UMoNbTiCr, which shows a BCC1 +BCC2 +IM phase.

It is necessary to explain how the Φ_s affect the formation of USP and UDP. As mentioned before, Φ_s is calculated by three terms, namely S_C, S_E and S_H. The former two are related to the configuration entropy and excess entropy, respectively, while the third term is related to the entropy from enthalpy, which is formulated as $S_H = |H_a|/T_m$ [40]. The H_a is calculated using the Miedema model, and the T_m is the average melting point, which can be estimated through the rule of mixing. This way, the higher the average melting point, the smaller is the S_H. In terms of USP and UDP equiatomic alloys in this work, they have similar ΔS_C and H_a, but different T_m. Accordingly, the higher the T_m, the lower is the S_H, thereby resulting in a larger numerator. However, we note that the S_E contribution is much stronger than the S_H due to the fact that the S_E value, as denominator, is at least one order lower than the numerator value (see details in Table S3). Our explanation points out that the Φ_s play a key role through the physical parameter T_m, which can be further explained as follows.

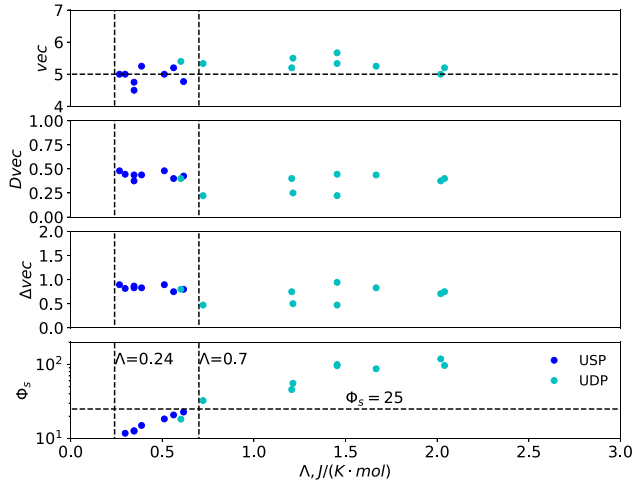

Fig. 4. *vec*, *Dvec*, Δvec and Φ_s against Λ of uranium-bearing DSSP (UDSSP). The USP and UDP denote the U-bearing single-phase and U-bearing dual-phase, respectively. Here it should be noted that we denote BCC+IM and BCC1 +BCC2 +IM as USP and UDP, respectively.

From the view point of solidification process, the as-cast microstructure is determined by synergy and competitive effect of thermal undercooling and consititutional undercooling. The effect of thermal undercooling on solidification is similar due to the same arc-melting casting method. The other is constitutional undercooling, caused by chemical composition, which leads to the formation of dendritic structures [43]. The stronger the constitutional undercooling effect, the larger the primary dendrite, and also the formation of dendrite branch. Therefore, the change of undercooling effect through tailing chemical composition is an important method to tune the microstructure, which is the main factors in this work.

In general, there exists element segregation in as-cast DSSAs with heavy element rich in dendrite arms and lighter elements rich in interdendritic regions. The difference in melting temperature between constituents encourages elements with higher T_m,such as Ta and Nb, to solidify first in the dendrite cores, ultimately forcing lighter elements with lower T_m, such as U, into the interdendritic region. The observed micro-segregation of the constituents is a manifestation of a non-equilibrium solidification of the alloys within the temperature range between the liquidus and solidus temperatures, so that homogeneous distribution of the alloying elements in the growing solid phases is kinetically restricted. The level of micro-segregation of the alloying elements is quantitatively described by a partition coefficient [44], $k = C_{da}/C_{idr}$, where C_{da} is the average concentrations in the centers of dendrite arms and C_{idr} is the average concentrations in interdendritic regions. The level of the element segregation increases with an expansion of the liquidus-solidus temperature range, which is generally associated with an increase in the difference in the melting temperature of the constituent elements, and an increase in the solidification rate. It has been evidenced that the segregation of an element *i* increases with an increase in the difference between the melting temperature and average melting temperature [44,45]. The U-rich interdendrite region ($k_U < 1.0$) and Ta- and Nb-rich dendrite arm ($k_{Ta} > 1.0$, and $k_{Nb} > 1.0$) in UMoNbTaTi have been confirmed [46]. Also, the lattice parameters can be estimated using the mixture rule of elements with errors within 5% [46]. Following this scenario, when the difference in lattice parameters of *idr* and *da* is small enough, resulting from the micro-segregation, it will still be single phase, however, when the discrepancy is large enough, the dual-phase will appear.

Although it seems that *vec* cannot distinguish the SP and MP, in fact *vec* do have contributions to the formation of USP and UDP. It has been suggested that the value of *vec* between 6.87 and 8.0 favors the formation of FCC and BCC duplex structures [24], which is evidenced in Fe-Co-Ni-Al system, while the *vec* value between 6.88 and 7.84 promotes the formation of σ phase for Cr- and/or V-containing high-entropy alloys [47]. Most recently, Li et al. [48] also confirms the effect of *vec* on the phase structures (see Fig. 3 in ref [48]): lower *vec* (2.8 ± 0.2) promote the hexagonal-closed packed (HCP) phase, while the larger value will make BCC (*vec*=5) and FCC (*vec*=8.5) phase more stable. These rules can be explained due to the fact that most of the FCC-based high-entropy alloys often use the elements with higher *vec* values, such as Fe, Co and Ni, while the BCC-based high-entropy alloys often use the elements with lower *vec* values, such as Ti, Nb and Mo. In Fig. 4, it shows that the *vec* has weak effect on the USP and UDP with no clearly boundary. This trend is also observed in other electron concentration related parameters.

There are several reasons for these inconsistent findings with previous results. First and foremost, we do not distinguish the HCP, BCC and FCC phases, and the BCC- and FCC-based phase are all denoted as SP phase. We also designate both the BCC1 +BCC2 phase and BCC+FCC phase as MP. Second, most of the elements used in our work are in regions of IV B, V B and VI B regions, which have similar *vec*. For USP and UDP alloys, they have a value between 5.0 and 6.0. These results clearly evidence that the third H − R rule, namely the electron concentration

rule, has weak effect on the formation of USP and UDP, but the Φ_s and Λ have a stronger effect. In contrast to previous findings [40], the UDP alloys show stronger entropy effect than USP alloys.

4. Conclusions

In this work, we have carried out a materials informatics investigation based on nearly 6000 multi-component alloys and 22 initial material descriptors to screen novel U-bearing DSSAs from 375 equiatomic multi-component alloys never known before. The 5 most important parameters, ΔS_c, Λ, Φ_s, γ, and $1/\Omega$, affecting DSSP formation are determined through the analyses of SHAP values. 190 out of 375 U-bearing alloys are predicted to be DSSAs. The experimental characterization of 20 randomly selected DSSAs show that (1) the predicted DSSAs alloys are in-line with the experimental findings with only three inconsistencies, (2) 5 most important features have stronger effect on the DSSAs formation ability, corresponding to the former two H − R rules, (3) a dual-phase microstructure with two BCC phases was predominantly found with a higher Φ_s ($\Phi_s > 25$) and λ value ($\Lambda > 0.7$), while a single BCC phase appeared with a lower Φ_s and λ value, (4) *vec*, the third H − R rule, has weak effect on the USP and UDP alloys. Moreover, we also substantiate how the ML models pick the candidate and in what way the material informatics can be related to the current understanding of formation mechanism of solid solution. We believe that the data-driven approach, such as that presented in this work, to screen material candidates, shed new light on the importance of entropy to the discovery of U-bearing DSSAs. A systematic experimental characterization on mechanical properties and microstructure of these DSSAs will be presented in the future.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Creative Development Foundation of China Academy of Engineering Physics (No. CX2019019), the National MCF Energy R&D Program (No. 2018YFE0312400) and the NSFC under the grant NO. U1930121. The Swedish Research Council, the Swedish Steel Producers' Association, the Swedish Foundation for Strategic Research, the Swedish Foundation for International Cooperation in Research and Higher Education, and the Hungarian Scientific Research Fund (research project OTKA 128229) are also acknowledged for financial support.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.mtcomm.2021.102960.

References

[1] S. Kaity, J. Banerjee, S. Parida, V. Bhasin, Structural, microstructural and thermal analysis of U-(6-x) Zr-xNb alloys (x= 0, 2, 4, 6), J. Nucl. Mater. 504 (2018) 234–250.
[2] D. Smirnova, A.Y. Kuksin, S. Starikov, Investigation of point defects diffusion in bcc uranium and U-Mo alloys, J. Nucl. Mater. 458 (2015) 304–311.
[3] N.W.S. Morais, D.A. Lopes, C.G. Schön, Effect of thermo-mechanical processing on microstructure and mechanical properties of U-Nb-Zr alloys: Part 2-U-3 wt% Nb-9 wt% Zr and U-9 wt% Nb-3 wt% Zr, J. Nucl. Mater. 502 (2018) 51–59.
[4] H.M. Volz, R.E. Hackenberg, A.M. Kelly, W. Hults, A. Lawson, R. Field, D. Teter, D. Thoma, X-ray diffraction analyses of aged U-Nb alloys, J. Alloy. Compd. 444 (2007) 217–225.
[5] Y.S. Kim, G. Hofman, J. Cheon, Recrystallization and fission-gas-bubble swelling of U-Mo fuel, J. Nucl. Mater. 436 (1–3) (2013) 14–22.
[6] K. Ghoshal, T. Kutty, S. Mishra, A. Kumar, Creep studies on U-7% Zr, U-7% Nb and U rich U-Nb-Zr alloys, J. Nucl. Mater. 432 (1–3) (2013) 20–22.
[7] K. Ghoshal, S. Kaity, S. Mishra, A. Kumar, Microstructural investigation of uranium rich U-Zr-Nb ternary alloy system, J. Nucl. Mater. 446 (1–3) (2014) 217–223.
[8] A. d.S. Ferreira, F.R. Longen, R.A.M. Gotardo, F.F. Ivashita, R. Barco, A. Paesano Júnior, Synthesis and structural characterization of U-Zr-Nb alloys, Mater. Res. 21 (1) (2018), e20170411.
[9] ⟨https://ltbridge.com/fuel-technology/metallic-fuel-technology/⟩.
[10] J.-W. Yeh, S.-K. Chen, S.-J. Lin, J.-Y. Gan, T.-S. Chin, T.-T. Shun, C.-H. Tsau, S.-Y. Chang, Nanostructured high-entropy alloys with multiple principal elements: novel alloy design concepts and outcomes, Adv. Eng. Mater. 6 (5) (2004) 299–303.
[11] B. Cantor, I. Chang, P. Knight, A. Vincent, Microstructural development in equiatomic multicomponent alloys, Mater. Sci. Eng. A 375 (2004) 213–218.
[12] F. Otto, Y. Yang, H. Bei, E.P. George, Relative effects of enthalpy and entropy on the phase stability of equiatomic high-entropy alloys, Acta Mater. 61 (7) (2013) 2628–2638.
[13] F. Zhang, S. Zhao, K. Jin, H. Bei, D. Popov, C. Park, J.C. Neuefeind, W.J. Weber, Y. Zhang, Pressure-induced fcc to hcp phase transition in Ni-based high entropy solid solution alloys, Appl. Phys. Lett. 110 (1) (2017), 011902.
[14] H. Huang, X. Li, Z. Dong, W. Li, S. Huang, D. Meng, X. Lai, T. Liu, S. Zhu, L. Vitos, Critical stress for twinning nucleation in CrCoNi-based medium and high entropy alloys, Acta Mater. 149 (2018) 388–396.
[15] Z. Li, S. Zhao, R.O. Ritchie, M.A. Meyers, Mechanical properties of high-entropy alloys with emphasis on face-centered cubic alloys, Prog. Mater. Sci. 102 (2019) 296–345.
[16] K. Jin, C. Lu, L. Wang, J. Qu, W. Weber, Y. Zhang, H. Bei, Effects of compositional complexity on the ion-irradiation induced swelling and hardening in Ni-containing equiatomic alloys, Scr. Mater. 119 (2016) 65–70.
[17] O. El-Atwani, N. Li, M. Li, A. Devaraj, J. Baldwin, M. Schneider, D. Sobieraj, J. Wróbel, D. Nguyen-Manh, S.A. Maloy, et al., Outstanding radiation resistance of tungsten-based high-entropy alloys, Sci. Adv. 5 (3) (2019), eaav2002.
[18] L. Ward, A. Agrawal, A. Choudhary, C. Wolverton, A general-purpose machine learning framework for predicting properties of inorganic materials, Npj. Comput. Mater. 2 (2016) 16028.
[19] F. Ren, L. Ward, T. Williams, K.J. Laws, C. Wolverton, J. Hattrick-Simpers, A. Mehta, Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments, Sci. Adv. 4 (4) (2018) eaaq1566.
[20] Y. Sun, H. Bai, M. Li, W. Wang, Machine learning approach for prediction and understanding of glass-forming ability, J. Phys, Chem. Lett. 8 (14) (2017) 3434–3439.
[21] R. Yuan, Z. Liu, P.V. Balachandran, D. Xue, Y. Zhou, X. Ding, J. Sun, D. Xue, T. Lookman, Accelerated discovery of large electrostrains in BaTiO3-based piezoelectrics using active learning, Adv. Mater. 30 (7) (2018), 1702884.
[22] C. Wen, Y. Zhang, C. Wang, D. Xue, Y. Bai, S. Antonov, L. Dai, T. Lookman, Y. Su, Machine learning assisted design of high entropy alloys with desired property, Acta Mater. 170 (2019) 109–117.
[23] D. Xue, P.V. Balachandran, J. Hogden, J. Theiler, D. Xue, T. Lookman, Accelerated search for materials with targeted properties by adaptive design, Nat. Commun. 7 (2016) 11241.
[24] S. Guo, C.T. Liu, Phase stability in high entropy alloys: formation of solid-solution phase or amorphous phase, Prog. Nat. Sci. -Mater. 21 (6) (2011) 433–446.
[25] X. Yang, Y. Zhang, Prediction of high-entropy stabilized solid-solution in multi-component alloys, Mater. Chem. Phys. 132 (2–3) (2012) 233–238.
[26] S. Guo, C. Ng, J. Lu, C. Liu, Effect of valence electron concentration on stability of fcc or bcc phase in high entropy alloys, J. Appl. Phys. 109 (10) (2011), 103505.
[27] Y. Zhang, T.T. Zuo, Z. Tang, M.C. Gao, K.A. Dahmen, P.K. Liaw, Z.P. Lu, Microstructures and properties of high-entropy alloys, Prog. Mater. Sci. 61 (2014) 1–93.
[28] Y. Zhang, Y.J. Zhou, J.P. Lin, G.L. Chen, P.K. Liaw, Solid-solution phase formation rules for multi-component alloys, Adv. Eng. Mater. 10 (6) (2008) 534–538.
[29] Y. Ye, Q. Wang, J. Lu, C. Liu, Y. Yang, High-entropy alloy: challenges and prospects, Mater. Today 19 (6) (2016) 349–362.
[30] M.C. Gao, C. Zhang, P. Gao, F. Zhang, L. Ouyang, M. Widom, J. Hawk, Thermodynamics of concentrated solid solution alloys, Curr. Opin. Solid St. M. 21 (5) (2017) 238–251.
[31] C.C. Tasan, Y. Deng, K.G. Pradeep, M. Yao, H. Springer, D. Raabe, Composition dependence of phase stability, deformation mechanisms, and mechanical properties of the CoCrFeMnNi high-entropy alloy system, Jom 66 (10) (2014) 1993–2001.
[32] Y. Tan, J. Li, Z. Tang, J. Wang, H. Kou, Design of high-entropy alloys with a single solid-solution phase: Average properties vs. their variances, J. Alloy. Compd. 742 (2018) 430–441.
[33] A.K. Singh, N. Kumar, A. Dwivedi, A. Subramaniam, A geometrical parameter for the formation of disordered solid solutions in multi-component alloys, Intermetallics 53 (2014) 112–119.
[34] G. Mansoori, N.F. Carnahan, K. Starling, T. Leland Jr, Equilibrium thermodynamic properties of the mixture of hard spheres, J. Chem. Phys. 54 (4) (1971) 1523–1525.
[35] Y. Ye, Q. Wang, J. Lu, C. Liu, Y. Yang, The generalized thermodynamic rule for phase selection in multicomponent alloys, Intermetallics 59 (2015) 75–80.
[36] G. Wang, L. Peng, K. Li, L. Zhu, J. Zhou, N. Miao, Z. Sun, Alkemie: An intelligent computational platform for accelerating materials discovery and design, Comput. Mater. Sci. 186 (2020), 110064.
[37] P. Jaccard, Distribution de la ore alpine dans la bassin de dranses et dans quelques regions voisines, Bull. Soc. Vaud. Sci. Nat. 37 (140) (1901) 241–272.
[38] S. Lundberg, S.I. Lee, A unified approach to interpreting model predictions (2017). arXiv:arXiv:1705.07874v2.

[39] M. Ali, PyCaret: An open source, low-code machine learning library in Python, pyCaret version 1.0.0 (April 2020).⟨https://www.pycaret.org⟩.
[40] Z. Wang, Y. Huang, Y. Yang, J. Wang, C.T. Liu, Atomic-size effect and solid solubility of multicomponent alloys, Scr. Mater. 94 (2015) 28–31.
[41] Y.F. Ye, Q. Wang, J. Lu, C.T. Liu, Y. Yang, Design of high entropy alloys: A single-parameter thermodynamic rule, Scr, Mater 104 (2015) 53–55.
[42] A. Takeuchi, K. Amiya, T. Wada, K. Yubuta, W. Zhang, A. Makino, Entropies in alloy design for high-entropy and bulk glassy alloys, Entropy 15 (9) (2013) 3810–3821.
[43] M.R. Rahul, S. Samal, G. Phanikumar, Metastable microstructures in the solidification of undercooled high entropy alloys, J. Alloy. Comp. 821 (2019), 153488.
[44] O.N. Senkov, G.B. Wilks, D.B. Miracle, C.P. Chuang, P. Liaw, Refractory high-entropy alloys, Intermetallics 18 (9) (2010) 1758–1765.
[45] H. Yao, J. Qiao, M. Gao, J. Hawk, S. Ma, H. Zhou, Y. Zhang, Nbtav-(ti,w) refractory high-entropy alloys: Experiments and modeling, Mater. Sci. Eng. A 674 (2016) 203–211.
[46] J. Shi, H. Huang, G. Hu, P. Zhang, C. Luo, Microstructure and mechanical properties of two uranium-containing high-entropy alloys, J. Alloy. Compd. 860 (1) (2020), 158295.
[47] M.H. Tsai, K.Y. Tsai, C.W. Tsai, C. Lee, C.C. Juan, J.W. Yeh, Criterion for sigma phase formation in cr- and v-containing high-entropy alloys, Mater. Res. Lett. 1 (2013) 207–212.
[48] R.X. Li, J.W. Qiao, P.K. Liaw, Y. Zhang, Preternatural hexagonal high-entropy alloys: a review, Acta Metall. Sin. 33 (2020) 1033–1045.