Available online at www.sciencedirect.com
ScienceDirect
Trans. Nonferrous Met. Soc. China 34(2024) 709−723
Transactions of Nonferrous Metals Society of China
www.tnmsc.cn

Accelerated design of Al−Zn−Mg−Cu alloys via machine learning

Yong-fei JUAN[1#], Guo-shuai NIU[2#], Yang YANG[2], Zi-han XU[1], Jian YANG[1], Wen-qi TANG[1], Hai-tao JIANG[1], Yan-feng HAN[1], Yong-bing DAI[1], Jiao ZHANG[1,3], Bao-de SUN[1,3]

1. Shanghai Key Lab of Advanced High-temperature Materials and Precision Forming and State Key Lab of Metal Matrix Composites, School of Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China;
2. Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China;
3. Collaborative Innovation Center for Advanced Ship and Deep-sea Exploration, Shanghai Jiao Tong University, Shanghai 200240, China

Received 21 September 2022; accepted 3 May 2023

Abstract: A machine learning-based alloy rapid design system (ARDS) was proposed to customize the preparation strategies for the desired properties or predict the alloy properties following the preparation strategies. For achieving this, three regression algorithms: linear regression (LR), support vector regression (SVR), and back propagation neural network (BPNN), were employed separately to train the multi-property prediction model, in which the machine learning (ML) model built using SVR was proved to be the best. Then, inspired by the generative adversarial network (GAN) algorithm, the ARDS was constructed. The predictive reliability of ARDS was examined, and for the accurate prediction of the preparation strategies, the upper limits of ultimate tensile strength (UTS), yield strength (YS), and elongation (EL) are about 790 MPa, 730 MPa, and 28%, respectively. Moreover, an ARDS-designed aluminum alloy with superior mechanical properties (764 MPa for UTS, 732 MPa for YS, and 10.1% for EL) was experimentally fabricated, further verifying the reliability of ARDS.

Key words: machine learning; alloy rapid design system; Al−Zn−Mg−Cu alloy; mechanical properties

1 Introduction

Al−Zn−Mg−Cu alloys, as multi-component aging alloys with excellent specific strength ratios, corrosion resistance, and low density, have been widely used in the aviation industry [1−3]. Currently, higher requirements are put forward on the properties of Al−Zn−Mg−Cu alloys to meet the development of the aviation industry, especially for strength and toughness. Fundamentally, the primary methods to enhance the performance of Al−Zn−Mg−Cu alloys mainly include the composition design, the control for heat treatment strategies, and the development of advanced processing techniques [4]. The emergence of some advanced processing technologies can certainly improve the performance of Al−Zn−Mg−Cu alloys, such as severe plastic deformation (SPD), which has realized the concurrent improvement in strength and plasticity [5]. However, these techniques are always defective in high costs, complex operations, and limited product sizes [6,7]. On the other hand, optimizing the material composition and establishing reasonable heat treatment strategies are two practical methods to improve the performance of the Al−Zn−Mg−Cu

# Yong-fei JUAN and Guo-shuai NIU contributed equally to this work
Corresponding author: Jiao ZHANG, Tel: +86-15801852318, E-mail: zj119@sjtu.edu.cn;
Yong-bing DAI, Tel: +86-13661593351, E-mail: ybdai@sjtu.edu.cn
DOI: 10.1016/S1003-6326(23)66429-5
1003-6326/© 2024 The Nonferrous Metals Society of China. Published by Elsevier Ltd & Science Press
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

alloys [8]. Nevertheless, the trialing scope of the unexplored element contents and heat treatment parameters is extensive for Al−Zn−Mg−Cu alloys, causing the routine trial- and-error method to be time-consuming and laborious. Following this, it must be meaningful to implement an intelligent design system for Al−Zn−Mg−Cu alloys.

Machine learning (ML) method, which can efficiently promote materials discovery based on plentiful algorithms and the strong ability in multi-dimensional modeling, has recently attracted considerable interest in materials research [9−12]. The current ML applications concerning the Al−Zn−Mg−Cu alloys mainly focus on the optimization of the heat treatment processes [13,14] and alloy properties, such as stress corrosion cracking resistance [15], ultimate tensile strength [16], and hardness [17]. However, the study on multi-property prediction is still scarce for Al−Zn−Mg−Cu alloys due to the complex and nonlinear relationship between the preparation strategies and material properties, especially for the strength and elongation, which are usually mutually exclusive [18,19]. Besides, different applications have different requirements on alloy properties. For example, the wing spar requires higher strength than the fuselage, and the latter requires higher corrosion resistance [1,2].

To address the above-mentioned limitations and facilitate the rational design of Al−Zn−Mg−Cu alloys, an alloy rapid design system (ARDS) for Al−Zn−Mg−Cu alloys was constructed in the present work based on the ML method, which can be used to customize the preparation strategies for the desired properties or predict the properties following the preparation strategies.

2 Data collection and processing

2.1 Data preparation

It is widely acknowledged that data quality and feature selection determine the upper limit of the ML model, and the ML models can only approach this limit. Abundant studies for Al−Zn−Mg−Cu alloys have been accumulated in recent decades, providing a data basis for applying ML methods. Herein, the preserved data were only Al−Zn−Mg−Cu alloys synthesized by traditional casting. Then, further requirement is that the details of alloy composition, processing technology, and material properties must be provided simultaneously in the studies. After preliminary screening, the dataset corresponding to the Al−Zn−Mg−Cu alloys, containing element contents, processing technology, and material properties, was successfully aggregated. The creation process of the dataset is illustrated in Fig. 1. The basic data were collected from 314 papers, whose sources are listed in Supplementary Materials.

Through the analysis of the alloy compositions in the retrieved literature, a total of 24 elements, i.e., Al, Zn, Mg, Cu, Zr, Si, Fe, Mn, Ti, Cr, Sc, La, Er, Y, Be, Pr, Yb, Ag, Ni, Ce, Li, Sr, Gd, and Sn, were selected as the initial composition features. Solution temperature and time, aging temperature and time were defined as the initial features of heat treatment parameters. For property features, in addition to the tensile strength and elongation, the other ones, such as yield strength, hardness, fracture toughness, and grain size, were also collected simultaneously. Next, the original dataset was further cleaned. The elements (Cu, Mg, Cr, Zn, Zr, Sc, Al) that appear

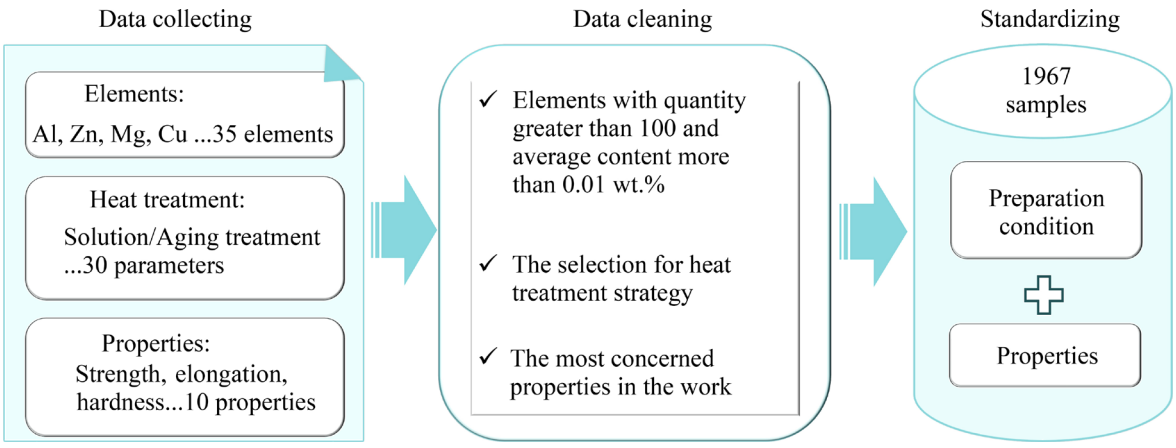

Fig. 1 Generation process of Al−Zn−Mg−Cu alloys dataset

more than 100 times with the average content being greater than 0.01 wt.% were selected as the final composition features. The single-step solution and aging were adopted as the heat treatment features since multi-stage heat treatment data are relatively few. The ultimate tensile strength, yield strength, and elongation, being the most concern to this work, were selected as property features. After verifying the reliability of the dataset, the standard scalar method was adopted to conduct the standardized processing of the dataset, and the features were normalized to [−1,1]. Then, the dataset containing 1967 labeled samples composed of the element contents, heat treatment parameters, and material properties was established.

2.2 Database analysis

Scatter plots for the relationships between the properties and the contents of major elements in the dataset are shown in Figs. 2 and 3, which indicate that the content of Zn has a positive effect on the ultimate tensile strength, while the effect is negative for the elongation. Figure 4 shows the relationships between the properties (ultimate tensile strength, yield strength and elongation), where ultimate tensile strength is positively correlated with yield strength but negatively with elongation. The relative distribution frequencies of the significant elements and properties are shown in Fig. 5, together with their Gaussian fitting. The content (wt.%) ratios of Zn/Mg, Mg/Cu, and (Zn+Cu)/Mg at the curve peaks are 2.5, 1.5, and 3.2, respectively, which are in good agreement with the other studies on Al−Zn−Mg−Cu alloys [20−22]. Meanwhile, the ranges between the upper and lower limits of the features are wide enough, as shown in Table 1 and Table 2, allowing the discovery of new materials with the desired properties.

3 Model development

In the present work, our goal is to develop an alloy rapid design method, which can be applied to customizing the preparation strategies for the desired properties or predict the properties following the preparation strategies.

3.1 Multi-property prediction model

Three ML algorithms, including linear regression (LR), support vector regression (SVR), and back propagation neural network (BPNN), were served to train and build the multi-property prediction

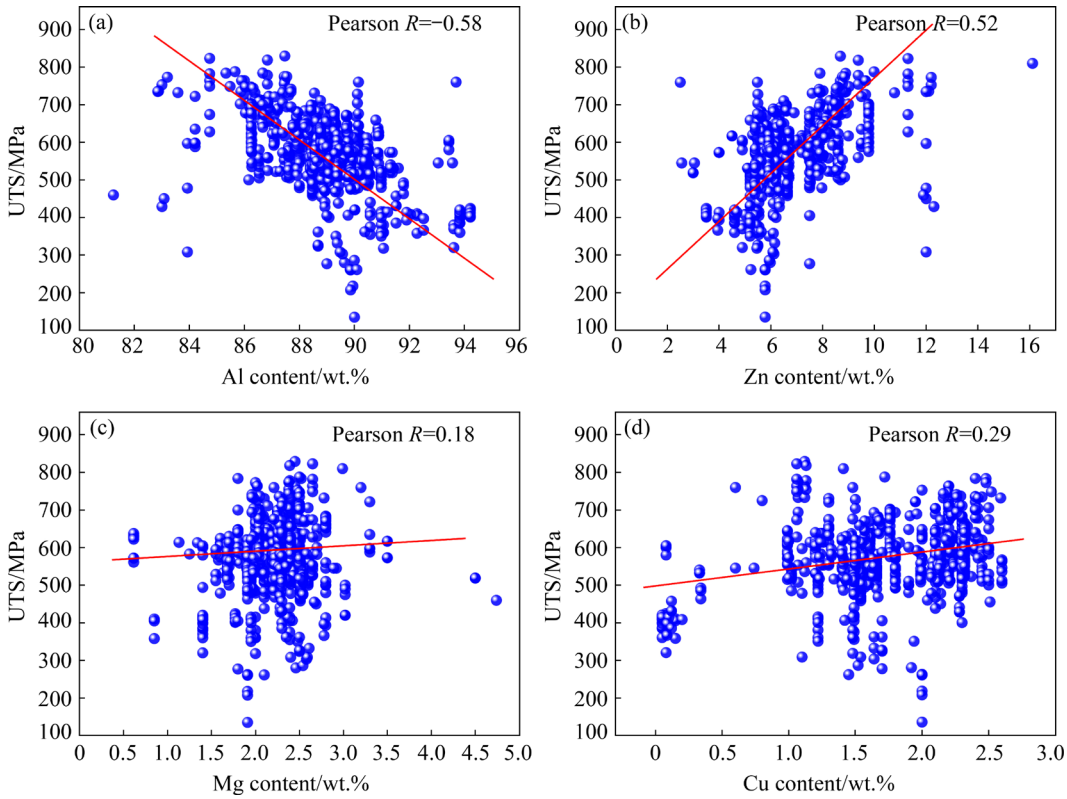

Fig. 2 Scatter plots showing relationships between UTS and Al (a), Zn (b), Mg (c), and Cu (d) contents (Pearson *R* refers to the correlation coefficient)

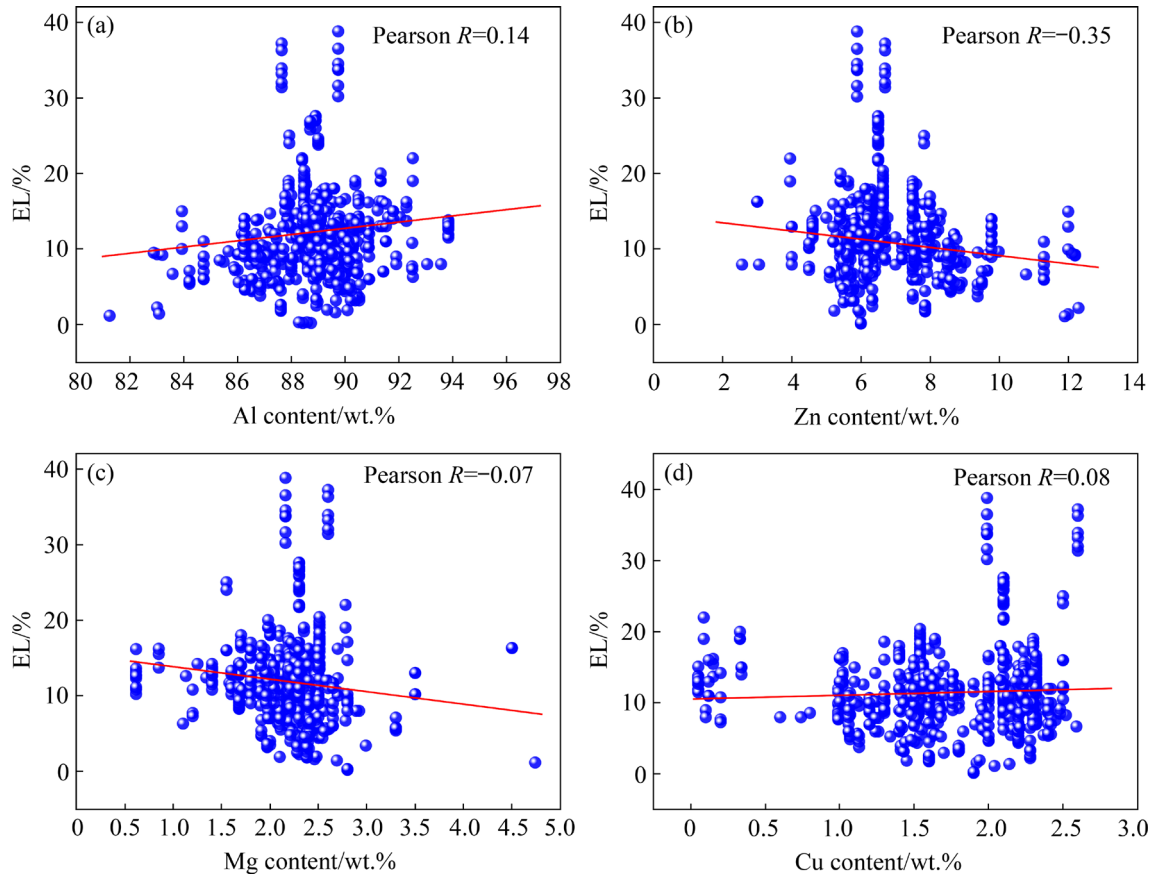

Fig. 3 Scatter plots showing relationships between EL and Al (a), Zn (b), Mg (c), and Cu (d) contents

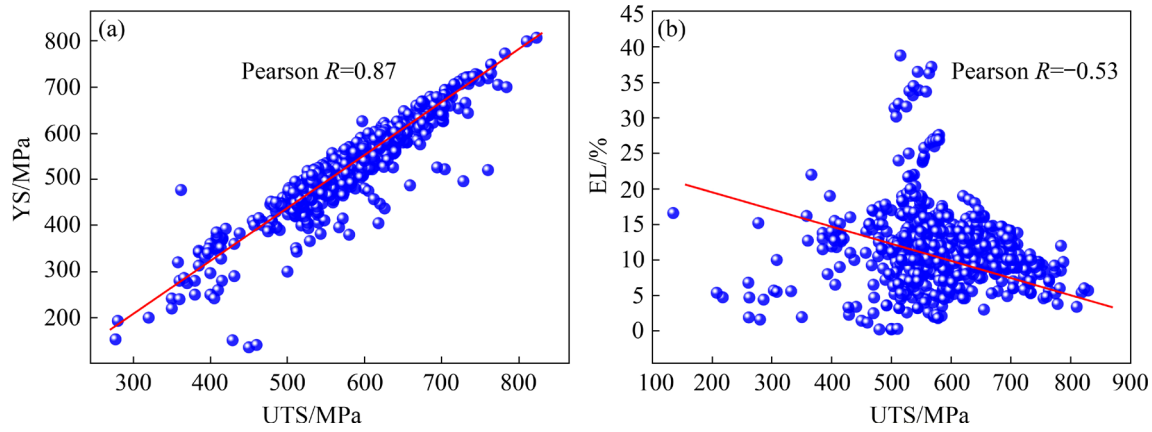

Fig. 4 Scatter plots showing relationships between UTS and YS (a), and UTS and EL (b)

Table 1 Analysis results of element contents and heat treatment parameters in dataset (T and t symbolize the temperature and time, respectively; subscripts ST and AT represent the solution and aging treatments, respectively; MIN, MAX and AVG indicate minimum, maximum and averaging values, respectively)

Item	Element content/wt.%						Heat treatment parameter			
	Al	Zn	Mg	Cu	Zr	Sc	T_{ST}/°C	t_{ST}/h	T_{AT}/°C	t_{AT}/h
MIN	80	2.5	0.6	0.1	0	0	420	0.2	25	0.1
MAX	94.2	16.1	4.7	2.7	0.3	0.6	520	48	240	190
AVG	89.2	6.3	2.2	1.7	0.1	0.1	470.7	2.2	121.8	23.2

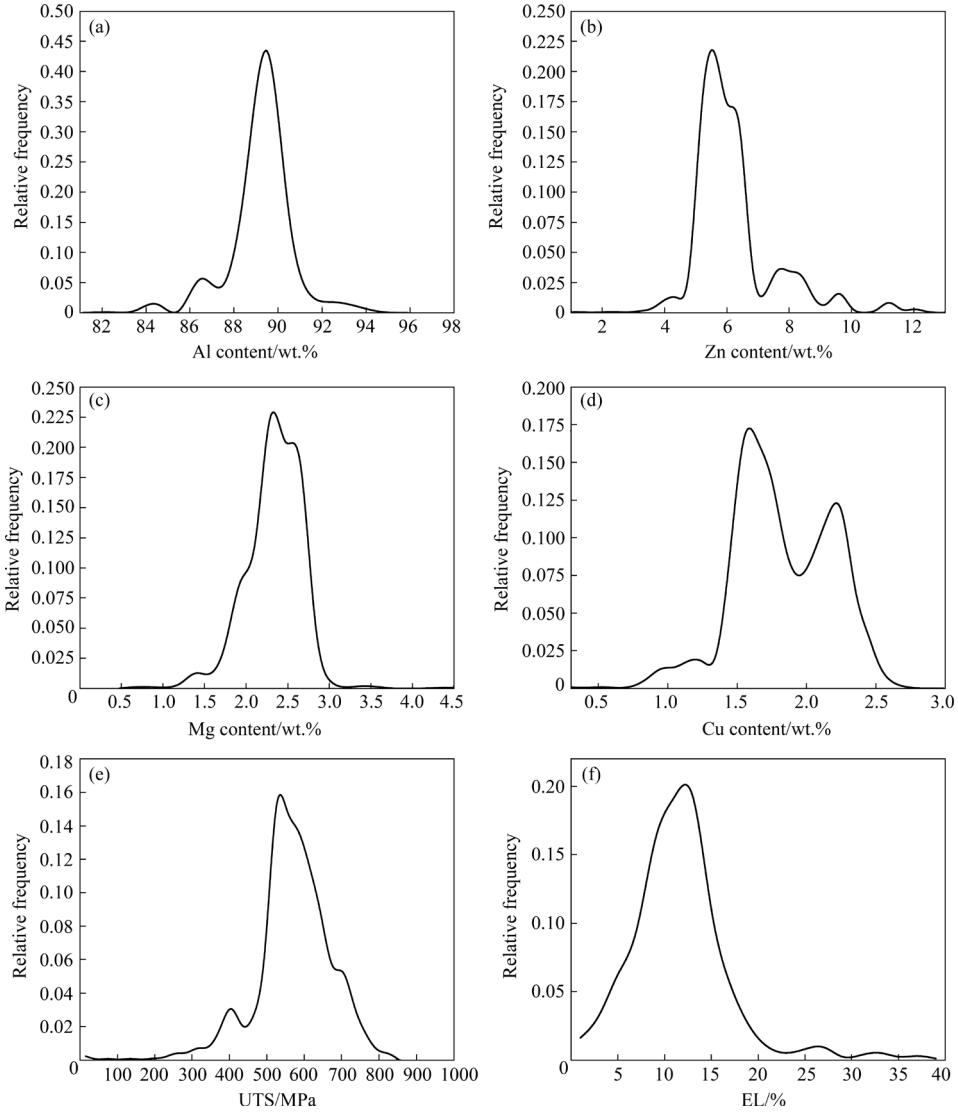

Fig. 5 Relative distribution frequencies of major element contents and properties: (a) Al; (b) Zn; (c) Mg; (d) Cu; (e) UTS; (f) EL

Table 2 Analysis results of main properties in dataset (H_V and σ represent the hardness and conductivity, respectively)

Item	UTS/MPa	YS/MPa	EL/%	H_V	σ/(S·m^{-1})
MIN	134.6	200	0.2	95	9
MAX	829.4	807	38.8	255.6	44.5
AVG	566.5	509.6	11.4	167.3	31

model. This model is named S-P, where S and P refer to the preparation strategies and material properties, respectively. The preparation strategies, which consisted of element contents and heat treatment parameters, are the input features of the model and material properties are the output ones. The reason for choosing LR, SVR and BPNN is that these three algorithms can solve the function problems in three different dimensions, namely linear function (LR), quadratic function (SVR) and high-dimensional function (BPNN), and show a progressive relationship. Firstly, a grid search was performed in this work to determine the best parameter for the SVR algorithm. Through the contrastive analysis, the ratio of training set to test one was set to be 7:3. Meanwhile, the regularization method was applied in the LR model to avoid over-fitting [23]. The optimization process for the BPNN algorithm was more complicated than those for SVR and LR. The reason was that the network structure in BPNN could be variable in the number of layers and nodes. We need to optimize BPNN with one or two hidden layers while the number of

nodes should be changed to reduce the deviations of BPNN; meanwhile, the iteration times and learning rate would significantly change in this process. Consequently, the hyperparameters such as network structure, iteration times, and learning rate all play crucial roles in the accuracy of the BPNN algorithm [24]. However, different from SVR and LR, there was no systematic theory for optimizing the BPNN. In this work, after repeated trial and verification, a three-layer neural network structure of 12-16-3 and a learning rate of 0.0035 were finalized to achieve high model performance.

As illustrated in Fig. 6, the predicted values via the three algorithms are compared with the actual values, and the deviations for both training and testing sets are listed in Table 3. It can be found that the relative errors and the deviations of SVR are the least. It is considered that SVR with the kernel function can construct data from the original space to higher dimensional feature space, sharing the advantages of dealing with scattered samples, nonlinearity, and high dimension, which meets our demands in this work. Following this, the S-P model was built based on the SVR algorithm.

3.2 Al−Zn−Mg−Cu alloy rapid design system

The three ML algorithms (LR, SVR, BPNN) were also employed to train and build the preparation-strategy prediction model, P-S model. However, the predicted results of P-S were far from reality, no matter which algorithm was adopted. There were even some negative values in the solution time and the element content of Cr. To solve this problem, we proposed an alloy rapid design system (ARDS), as described in Fig. 7, by referring to the idea of the generative adversarial

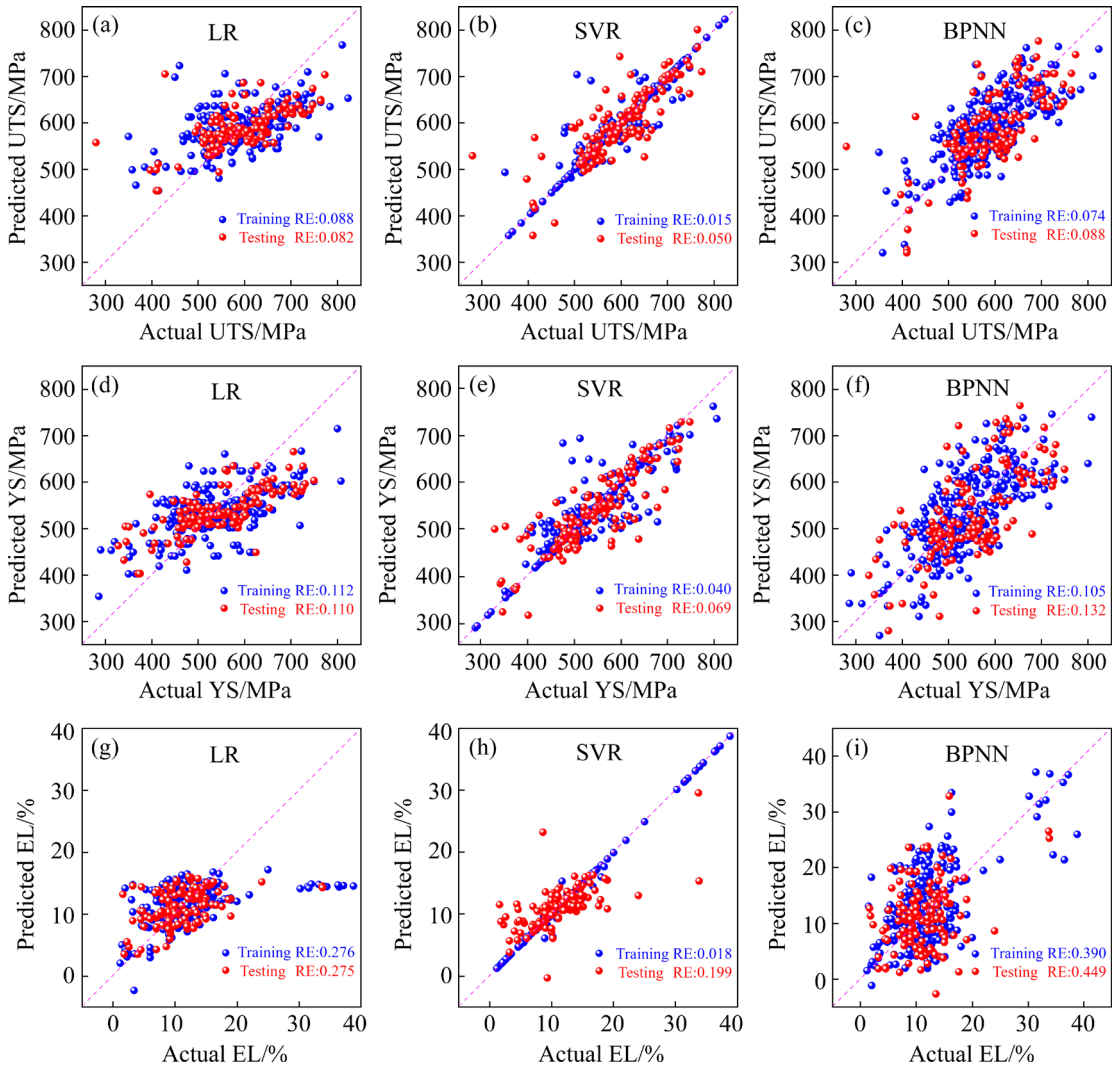

Fig. 6 Predicted results of three ML models compared with actual ones: (a, b, c) UTS; (d, e, f) YS; (g, h, i) EL (RE represents the relative error)

network (GAN) algorithm [25]. GAN has two main parts, including the generative and the discriminative models. The task of the generative model is to generate instances similar to the original data. The discriminant model is employed to judge whether the given instance is trustworthy and reliable. Therefore, the principle of the GAN algorithm is in line with our original intention for alloy design. WANG et al [26] reported a design strategy for copper alloys, whose inspiration also seems to come from the GAN algorithm. In our system, the preparation strategies were predicted firstly by the P-S model based on the target properties, and then the strategies were input into S-P model. Afterward, the predictions of the S-P model were compared with the target properties, and the deviation for each property was calculated by the formula: $|A-B|/B$, where A is the prediction value and B is the target value. The preparation strategy was output as Result-1 when the deviation for every property was less than a predefined threshold (PT), which was selected to ensure that the results were accurate. Otherwise, the ARDS system would continue trial and error.

The prediction system is almost impossible to succeed when the target properties deviate too much from those appeared in the dataset. Thus, the largest number of training times (LNTT) should be set to terminate otherwise endless trialing. If any of the LNTT training cycles, there is always a deviation for any of the properties bigger than PT, then the preparation strategy with the summed deviations for properties being minimum through LNTT cycles was output as Result-2. In this work, the value of PT was set to be 0.05 after a careful examination, which is appropriate for ARDS. The value of LNTT was set as follows: Firstly, the dataset was split into training and test sets with six different ratios, i.e., 9:1, 8:2, 7:3, 6:4, 5:5, and 4:6. Then, for each split way, the training set was randomly chosen from the dataset for 50 times. While one training for the whole dataset was also included, an LNTT value of 301 was obtained. This value was found to be a relatively good choice concerning the predictive ability and efficiency after trying other LNTT values, e.g., 241, 361 and 421. At last, the system was checked with its P-S model built by different algorithms. It was found that the system performed best in both efficiency and accuracy when the P-S model was built by the SVR algorithm. For instance, when target properties were set as follows: ultimate tensile strength of 700 MPa, yield strength of 650 MPa and elongation of 10%, the system output the preparation strategy as Result-1 by only one training cycle if SVR algorithm was applied, but by more than seven training cycles if BPNN or LR algorithm was applied. So, both S-P and P-S models in the ARDS were constructed based on the SVR algorithm. Compared with the work of WANG et al [26] the

Table 3 Deviations of training and test sets for three ML models (Train-error and Test-error represent deviations of training and test sets, respectively)

Model	Train-error of UTS	Test-error of UTS	Train-error of YS	Test-error of YS	Train-error of EL	Test-error of EL
LR	0.0810	0.0824	0.1007	0.1047	0.2409	0.2582
SVR	0.0125	0.0569	0.0457	0.0626	0.0428	0.1931
BPNN	0.0586	0.0683	0.0754	0.0781	0.2385	0.2585

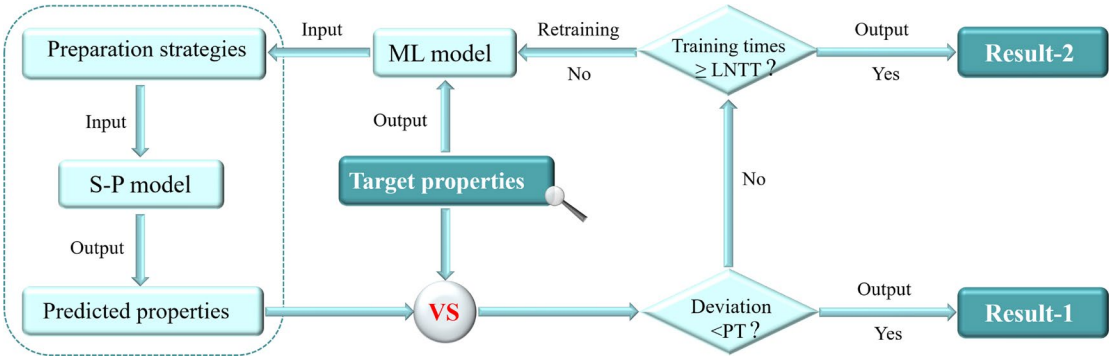

Fig. 7 Running steps of ARDS (The system outputs two types of results, i.e., Result-1 and Result-2)

difference of ARDS lies in the addition of heat treatment parameters in the preparation strategy and the discussion on more algorithms. Besides, the video demonstrations of the prediction processes are attached as supplemental materials (see "Operation interface for S-P" and "Operation interface for P-S" in the Supplemental Materials).

3.3 Practicability and validity of ARDS

In order to check whether the ARDS has achieved the design purpose, we compared the prediction results with the experimental data from the literature.

Firstly, the predictive ability for material properties was checked, where only the leftmost part enclosed by the dashed line in Fig. 7 was employed. Based on the analysis of the dataset and considering that Zn is the critical element affecting the performance of Al−Zn−Mg−Cu alloys [27,28], six groups of prediction schemes (Gropes 1#−6# in Table 4) were designed following the Zn content increasing from 6.0 wt.% to 8.5 wt.% with an interval of 0.5 wt.%. The properties predicted according to the six groups of preparation strategies were compared with the experimental data outside the dataset [29−34], as described in Table 4. It is clear that the predicted results are in good agreement with the actual values.

The ability of the system to predict the preparation strategies was also tested with properties as the input. The range of tensile strength, one of the most important properties of Al−Zn−Mg−Cu alloys [35−37], was set to be 580−780 MPa with an interval of 40 MPa. The values of the corresponding yield strength and elongation were designated following the linear relationships shown in Fig. 4. Based on this, six groups of properties were determined, as shown in Table 5, the average values obtained in five predictions were taken as the final results, where the preparation strategies predicted by the system are generally in line with the experimental data [38−43], thus verifying the practicability and validity of the design system again.

Though the predictions shown in Table 5 may be encouraging, it should be pointed out that, for Groups 1#−5#, all the predictions were those output as Result-1, but for Group 6#, i.e., the ultimate tensile strength increased to 780 MPa, among five predictions, two were output as Result-2. This observation implies that the prediction range of the system is not boundless, as can be deduced from the scarcity of data with ultimate tensile strength higher than 800 MPa in the original dataset (see Fig. 5(e)). Obviously, for the system, the upper limits of properties for achieving accurate predictions of the corresponding preparation strategies do exist. By utilizing experimental data outside the dataset, including some newly collected [44−48] and some shown in Table 5, the upper limits were checked in

Table 4 Comparison of properties predicted according to six preparation strategies with experimental data (The Zn content is changed from 6 wt.% to 8.5 wt.% with an interval of 0.5 wt.%)

Group No.	Element content/wt.%						Heat treatment parameter				Property (output)		
	Cu	Mg	Cr	Zn	Zr	Sc	T_{ST}/°C	t_{ST}/h	T_{AT}/°C	t_{AT}/h	UTS/MPa	YS/MPa	EL/%
1#	2.00	2.00	0	6.00	0	0	470.0	1.0	120.0	24.0	557.0	514.0	16.4
Ref. [29]	2.12	2.20	0	6.06	0.11	0	473.0	1.0	120.0	24.0	562.0	525.0	−
2#	2.00	2.50	0	6.50	0	0	470.0	1.0	120.0	24.0	594.6	570.7	13.3
Ref. [30]	2.20	2.35	0	6.40	0	0	475.0	0.5	120.0	24.0	610.2	534.8	15.9
3#	2.00	2.00	0	7.00	0	0	470.0	2.0	120.0	24.0	622.6	591.3	11.2
Ref. [31]	1.20	2.00	0.1	7.20	0.10	0	470.0	2.0	120.0	24.0	613.5	571.2	10.7
4#	2.00	2.50	0	7.50	0.10	0	470.0	2.0	120.0	24.0	651.2	613.4	9.8
Ref. [32]	1.24	2.20	0	7.60	0.13	0	470.0	2.0	120.0	24.0	638.0	556.0	10.2
5#	2.00	2.50	0	8.00	0.10	0.11	470.0	2.0	120.0	24.0	695.1	655.7	10.8
Ref. [33]	2.42	1.98	0	7.91	0.12	0	475.0	1.0	120.0	24.0	704.0	660.0	9.7
6#	2.00	2.50	0	8.50	0.11	0.12	475.0	2.0	120.0	24.0	727.8	694.0	10.1
Ref. [34]	2.16	2.38	0.17	8.53	0.16	0	480.0	2.0	130.0	24.0	743.9	728.1	9.2

Table 5 Comparison between preparation strategies predicted by ARDS and published experimental data (The ultimate tensile strength is changed from 580 to 780 MPa with an interval of 40 MPa)

Grope No.	Property (input)			Element content/wt.%						Heat treatment parameter			
	UTS/MPa	YS/MPa	EL/%	Cu	Mg	Cr	Zn	Zr	Sc	T_{ST}/°C	t_{ST}/h	T_{AT}/°C	t_{AT}/h
1#	580	540	13	1.74	2.07	0.01	6.23	0.11	0	470.1	1.5	119.8	23.8
Ref. [38]	583	502	12	1.5	2.58	0.19	6	0	0	480	1	120	24
2#	620	580	12	1.84	2.08	0	6.79	0.1	0.01	470.5	1.9	120	23.9
Ref. [39]	616	553.5	12.7	1.53	1.68	0	7.55	0.11	0	470	2	135	20
3#	660	620	11	1.85	2.25	0.01	7.81	0.09	0.01	470	2	120	23.9
Ref. [40]	643.2	607.2	14.6	2.1	1.8	0	8	0.18	0	470	1	121	24
4#	700	660	10	2.01	2.29	0.01	8.19	0.09	0.09	472.6	2.1	120.1	24
Ref. [41]	709.7	683.6	8.9	2.2	2.5	0	8.6	0.16	0	480	2	130	24
5#	740	700	9	2.16	2.2	0	8.74	0.11	0.1	473	2.1	120.5	24.5
Ref. [42]	734	644	9.8	2.26	1.96	0	7.87	0.1	0.11	480	1	120	24
6#	780	740	8	2.2	2.31	0	9.18	0.13	0.22	474.2	2	120	24.2
Ref. [43]	783.9		8.5	2.48	2.52	0	8.89	0.17	0.61	455	2	120	24

this work. The experimental properties were input as the target ones (see Fig. 7), then the corresponding preparation strategies were predicted via the system. The relative deviation (Δd) between the predicted and experimental preparation strategies was calculated as follows:

$$\Delta d=\left|\frac{T_{ST}^{p}-T_{ST}^{e}}{T_{ST}^{e}}\right|+\left|\frac{t_{ST}^{p}-t_{ST}^{e}}{t_{ST}^{e}}\right|+\left|\frac{T_{AT}^{p}-T_{AT}^{e}}{T_{AT}^{e}}\right|+\left|\frac{t_{AT}^{p}-t_{AT}^{e}}{t_{AT}^{e}}\right|+\sum_{i=1}^{n}\left|\frac{c_i^{p}-c_i^{e}}{c_i^{e}}\right| \quad (1)$$

where superscripts p and e refer to the prediction and experiment, respectively; n stands for the number of the critical elements for Al−Zn−Mg−Cu alloys; c denotes the element content.

Five predictions were performed as before, and the averaged results $\overline{\Delta d}$ are presented in Fig. 8. As can be found from Fig. 8, for any of the properties, with an increasing value of it, $\overline{\Delta d}$ firstly fluctuates around a relatively small value, and after a critical value of the property, it rises quickly. At the same time, the circle shown in the figure changes from filled to partially filled one, demonstrating that more and more predictions are those output as Result-2. It can be concluded from Fig. 8 that for the accurate prediction of the preparation strategies, the upper limits (i.e., the critical values) of the ultimate tensile strength, yield strength, and elongation are about 790 MPa, 730 MPa, and 28%, respectively.

4 Experimental verification

Herein, the input items of the P-S model in ARDS are fewer than the output ones; that is, the prediction difficulty for the preparation strategy is significantly greater than that for the alloy property. Meanwhile, we found that the improvement of the property of Al−Zn−Mg−Cu alloys is accompanied by the increased Sc content, as evident from Table 5. Therefore, we chose the group of properties with the highest UTS, i.e., Grope 6# in Table 5, for experimental verification, where the input items were 780.0 MPa (UTS), 740.0 MPa (YS), 8.0% (EL); and the design results were as follows: Al−9.18Zn−2.31Mg−2.20Cu−0.13Zr−0.22Sc (named ARDS-1 alloy), and the solution and aging parameters were 474.2 °C for 2.0 h and 120.0 °C for 24.2 h, respectively. Besides ARDS-1 alloy, we conducted a controlled trial in the experimental verification, where the alloy composition was set to Al−9.18Zn− 2.31Mg−2.20Cu−0.13Zr, and the other process parameters were the same as ARDS-1, to estimate the reliability of the predicted results and the influence of Sc element on the properties of Al−Zn−Mg−Cu alloys.

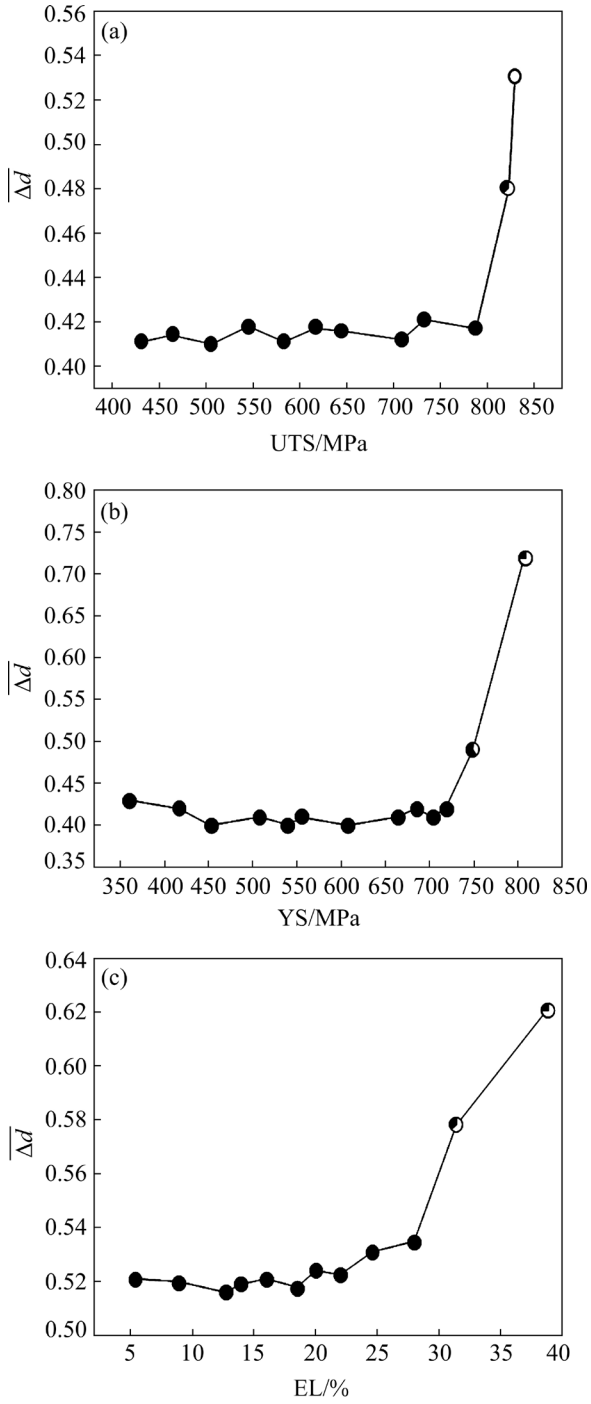

Fig. 8 Averaged relative deviation ($\overline{\Delta d}$) between predicted and experimental preparation strategies for different target properties (The filled or empty (●, ○) circle indicates that all five predictions were those output as Result-1 or Result-2. The partially filled circle (◐) denotes that, among five predictions, some were output as Result-1, while the others were output as Result-2, and the relative areas of the filled and empty parts of a circle demonstrate the ratio between the number of predictions output as Result-1 to that output as Result-2.): (a) UTS; (b) YS; (c) EL

4.1 Experimental procedures

The two group alloys were both smelted and cast into ingots of approximately 3 kg, where the raw materials were pure metals (Al, Mg, and Zn) with 99.9 wt.% purity and master alloys (Al−50wt.%Cu, Al−10wt.%Zr, and Al−20wt.%Sc). The as-cast alloys were treated by a three-stage homogenization: 418 °C for 5 h, 465 °C for 24 h, and 473 °C for 24 h in resistance furnace, followed by a multistage rolling strategy, including fifteen-stage hot rolling and nine-stage cold rolling. Then, the predicted results for the heat treatment system, i.e., 474.2 °C for ST-*T*, 2.0 h for ST-*t*, 120.0 °C for AT-*T*, and 24.2 h for AT-*t*, were used for the subsequent solution and aging treatments.

4.2 Experimental results

The as-cast ARDS-1 presents a fully fine equiaxed microstructure, as shown in Figs. 9(d−f), rather than the coarse dendrites that can be discovered in the control group, as shown in Figs. 9(a−c). Calculated by ImageJ software, the average grain size of ARDS-1 was about 40 μm, while it was over 100 μm for the control group, indicating that adding Sc can significantly refine the grain. The area scan results for the as-cast ARDS-1 are displayed in Figs. 9(g−l), from which a relatively homogeneous distribution of the precipitated Al_3(Sc,Zr) particles in the Al matrix can be seen. The line scan results in Fig. 9(f) further illustrate the microstructure characteristics of Al_3(Sc,Zr) particles, which show a square and regular core−shell structure. Compared with Sc, the distribution of the Zr element in particles is more concentrated in the core. The difference between ARDS-1 and the control group is whether the Sc element is added. Hence, the experimental results confirm that Sc plays a significant role in grain refinement, and it can combine with Zr to form the nucleation core, i.e., Al_3(Sc,Zr) particles.

In order to reveal the mechanism behind different strengthening effects of ARDS-1 alloy, the microstructures of ARDS-1 aged at 120 °C for 24.2 h were observed. The incident beam is parallel to ⟨011⟩ direction in the α(Al) matrix (B=⟨011⟩$_\alpha$). Figures 10(a) and (b) present a wealth of precipitated phases with various sizes in ARDS-1 alloy; the calculated results indicate that the phases within the grain are 2−20 nm in size, and these nano

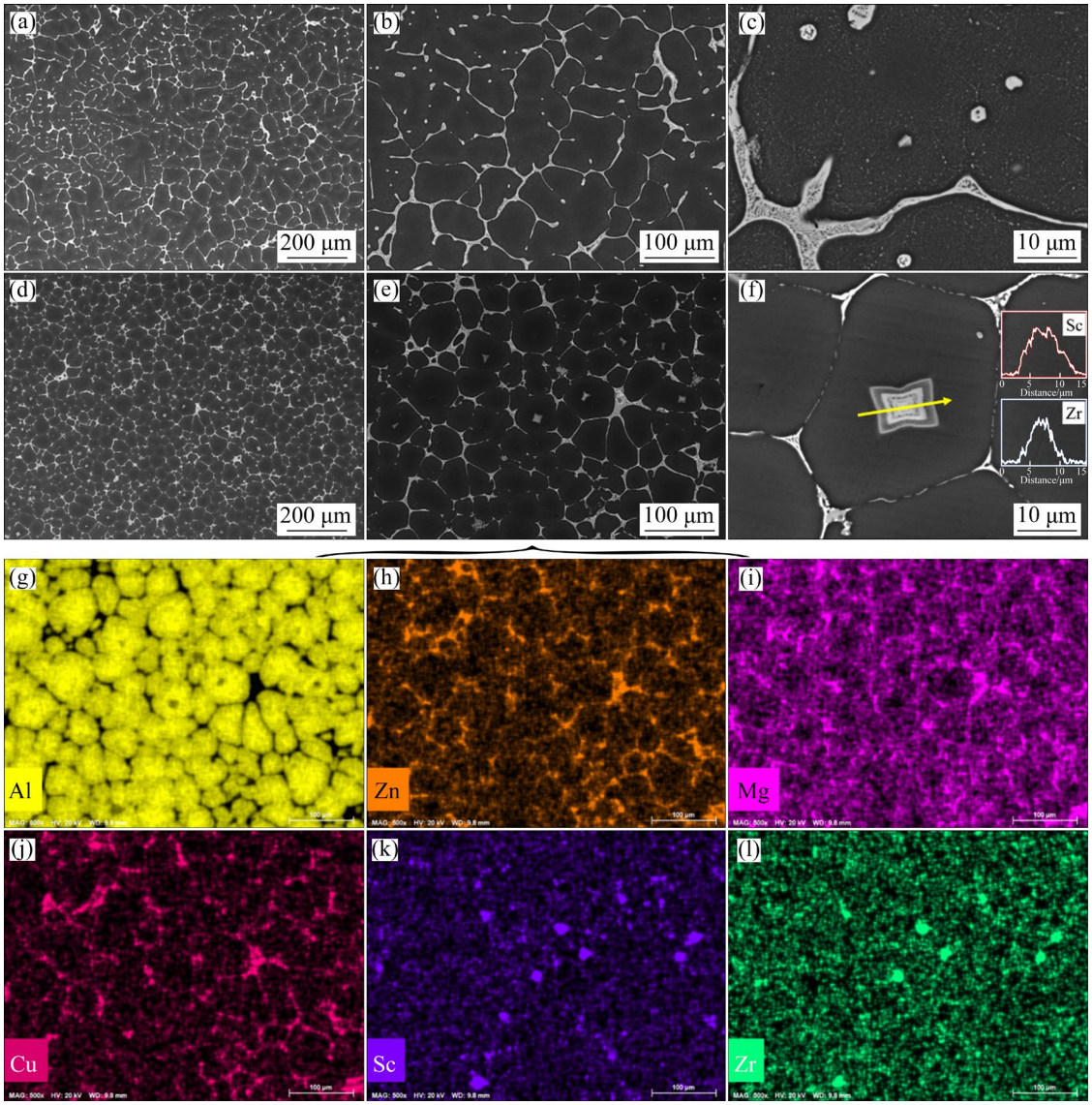

Fig. 9 Microstructures of control group (a−c), ARDS-1 (d−f) and area scan results (g−l) corresponding to microstructure image in (e) (The line scan results in (f) show the microstructure characteristics of $Al_3(Sc,Zr)$ particles)

precipitates play a significant role in precipitation strengthening and are decisive in enhancing the alloy properties. On the other hand, the interfacial energy at the grain boundary is usually high, which can provide more nucleation energy for the heterogeneous precipitation of $MgZn_2$ phases. It is obvious from Figs. 10(b−d) that some coarse $MgZn_2$ phases are found at the grain boundary. As shown in Fig. 10(d), a relatively homogeneous distribution of the precipitated $Al_3(Sc,Zr)$ particles in the Al matrix is observed. These fine particles effectively impede the motion of dislocations, which can disorderly intertwine to form dislocation pile-ups, thereby increasing the recrystallization temperature of the alloy and generating substructure strengthening. Furthermore, $Al_3(Sc,Zr)$ can also stabilize the vacancy and weaken the diffusion ability of solute atoms, thus reducing the vacancy and solute concentration gradients from the grain boundary to the grain interior, and narrowing down the width of the grain boundary precipitation-free zone (PFZ).

To further verify the reliability of ARDS, the mechanical properties of the designed alloy ARDS-1 and the control group Al−9.18Zn−2.31Mg−2.20Cu−0.13Zr were tested, and the tensile stress−strain curves of the two alloys are shown in Fig. 11(a). The comparison of comprehensive properties between them is evident in Fig. 11(b), where the measured values of UTS, YS and EL for designed ARDS-1 alloy/control group are

764/690 MPa, 732/663 MPa, and 10.1%/6.0%, respectively. The average prediction error between the experimental results and predicted ones (780 MPa for UTS, 740 MPa for YS, and 8.0% for EL) is 9.8%, which is close to those from similar studies [16,49,50]. This demonstrates that the experimental results agree with the predicted values, further verifying the reliability of ARDS in designing aluminum alloys. It is noteworthy that the only difference between the two alloys is whether they contain the Sc element, indicating that the addition of Sc plays a decisive role in improving the properties of the alloy, which echoes the analysis of the microstructures.

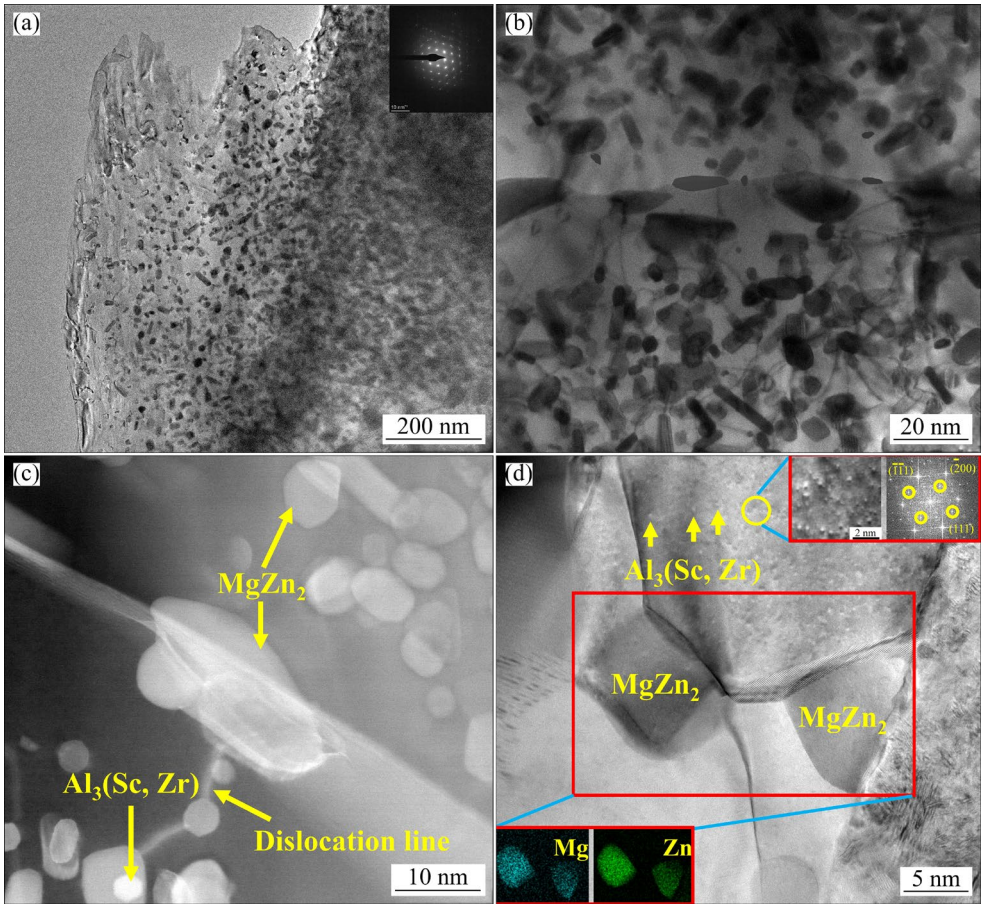

Fig. 10 Microstructures of ARDS-1 alloy: (a, b) TEM images; (c) STEM image showing nanoscale precipitated phases; (d) HAADF image of $Al_3(Sc,Zr)$ particles and corresponding fast-Fourier-transform (FFT) diffractogram, and elemental mapping of $MgZn_2$

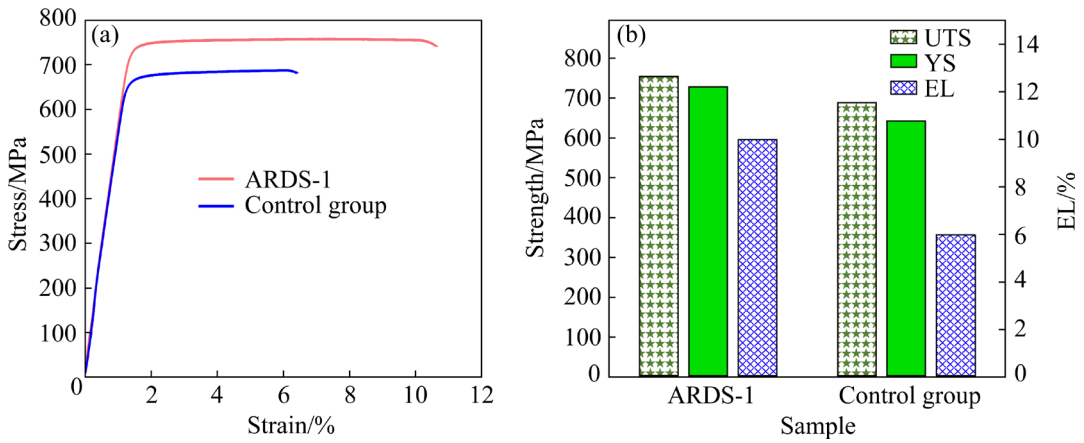

Fig. 11 Stress−strain curves (a) and comparison of comprehensive properties (b) between ARDS-1 and control group

5 Conclusions

(1) A dataset for Al−Zn−Mg−Cu alloys composed of three parts: element contents, heat treatment parameters, and material properties, was established by collecting the experimental results from relevant references.

(2) Based on the multi-property prediction model that was constructed based on the SVR algorithm and drawing inspiration from the GAN algorithm, an alloy rapid design system (ARDS) was constructed, which can be used to customize the preparation strategies for the desired properties or predict the properties following the preparation strategies

(3) The practicability and predictive reliability of this system were examined using published data. For the accurate prediction of the preparation strategies, the upper limits of the UTS, YS, and EL are about 790 MPa, 730 MPa, and 28%, respectively.

(4) An alloy designed by ARDS (ARDS-1) with remarkable comprehensive properties: 764 MPa of UTS, 732 MPa of YS, and 10.1% of EL, was experimentally fabricated, and the analysis results indicated that the addition of Sc plays a decisive role in improving the properties of Al−Zn−Mg−Cu alloys.

CRediT authorship contribution statement

Yong-fei JUAN: Conceptualization, Methodology, Investigation, Data collection, Formal analysis, Writing − Original draft, Software; **Guo-shuai NIU:** Formal analysis, Methodology, Software; **Yang YANG:** Formal analysis, Validation, Methodology; **Zi-han XU:** Data collection; **Jian YANG:** Formal analysis; **Wen-qi TANG:** Data collection; **Hai-tao JIANG:** Data collection; **Yan-feng HAN:** Formal analysis, Project administration; **Yong-bing DAI:** Conceptualization, Methodology, Formal analysis, Writing − Review & editing; **Jiao ZHANG:** Formal analysis, Funding acquisition, Supervision; **Bao-de SUN:** Funding acquisition, Supervision.

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.

Data availability

The data that support the results reported in this paper and other findings of this work are available from the corresponding author upon reasonable request.

Acknowledgments

This work was financially supported by the National Key Research and Development Program of China (No. 2020YFB0311201), and the National Natural Science Foundation of China (No. 51627802).

Supplementary Materials

Supplementary Material in this paper can be found at: http://tnmsc.csu.edu.cn/download/01-p0709-2022-1058-Supplemetary_materials.pdf.

References

[1] POLLOCK T M. Alloy design for aircraft engines [J]. Nature Materials, 2016, 15: 809−815.

[2] WILLIAMS J C, STARKE E A. Progress in structural materials for aerospace systems [J]. Acta Materialia, 2003, 51: 5775−5799.

[3] PARK S Y, KIM W J. Difference in the hot compressive behavior and processing maps between the as-cast and homogenized Al−Zn−Mg−Cu (7075) alloys [J]. Journal of Materials Science & Technology, 2016, 32: 660−670.

[4] YUAN Liang-liang, GUO Ming-xing, YAN Yong, FENG Wei-jun, LIU Zan-yang, ZHUANG Lin-zhong. Theoretical design and distribution control of precipitates and solute elements in Al−Zn−Mg−Cu alloys with heterostructure [J]. Transactions of Nonferrous Metals Society of China, 2021, 31: 3328−3341.

[5] ZHANG Yi-dong, JIN Shen-bao, PATRICK W T, LIAO Xiao-zhou, MAXIM Y M, RUSLAN Z V, LIU Ji-zi, JULIE M C, SIMON P, SHA Gang. Dynamic precipitation, segregation and strengthening of an Al−ZnMg−Cu alloy (AA7075) processed by high-pressure torsion [J]. Acta Materialia, 2019, 162: 19−32.

[6] SRIVASTAVA A K, SRIVSTAVA V C, GLOTER A, OJHA S N. Microstructural features induced by spray processing and hot extrusion of an Al−18%Si−5%Fe−1.5%Cu alloy [J]. Acta Materialia, 2006, 54: 1741−1748.

[7] SABBAGHIANRAD S, LANGDON T G. A critical evaluation of the processing of an aluminum 7075 alloy using a combination of ECAP and HPT [J]. Materials Science and Engineering: A, 2014, 596: 52−58.

[8] XIAO Hong-yu, LI Yu-gang, GENG Ji-wei, LI Hong-ping, WANG Ming-liang, CHEN Dong, LI Zhu-guo, WANG Hao-wei. Effects of nano-sized TiB_2 particles and Al_3Zr dispersoids on microstructure and mechanical properties of Al−Zn−Mg−Cu based materials [J]. Transactions of Nonferrous Metals Society of China, 2021, 31: 2189−2207.

[9] JUAN Yong-fei, DAI Yong-bing, YANG Yang, ZHANG Jiao. Accelerating materials discovery using machine learning [J]. Journal of Materials Science & Technology, 2021, 79: 178−190.

[10] RACCUGLIA P, ELBERT K C, ADLER P D F, FALK C, WENNY M B, MOLLO A, ZELLER M, FRIEDLER S A, SCHRIER J, NORQUIST A J. Machine-learning-assisted materials discovery using failed experiments [J]. Nature, 2016, 533: 73−76.

[11] LIU Yue, WU Jun-ming, WANG Zhi-chao, LU Xiao-Gang, MAXIM A, SHI Si-qi, WANG Chong-yu, YU Tao. Predicting creep rupture life of Ni-based single crystal superalloys using divide-and-conquer approach based machine learning [J]. Acta Materialia, 2020, 195: 454−467.

[12] JUAN Yong-fei, NIU Guo-shuai, YANG Yang, DAI Yong-bing, ZHANG Jiao, SUN Bao-de. Knowledge-aware design of high-strength aviation aluminum alloys via machine learning [J]. Journal of Materials Research and Technology, 2023, 24: 346−361.

[13] FANG S F, WANG M P, SONG M. An approach for the aging process optimization of Al−Zn−Mg−Cu series alloys [J]. Materials & Design, 2009, 30: 2460−2467.

[14] VAHID A, RANA S, GUPTA S, VELLANKI P, DORIN T. New bayesian-optimization-based design of high-strength 7xxx-series alloys from recycled aluminum [J]. JOM, 2018, 70: 2704−2709.

[15] CAO Xin-yu, ZHANG Ying-bo, LI Jia-heng, CHEN Hui. Composition design of 7xxx aluminum alloys optimizing stress corrosion cracking resistance using machine learning [J]. Materials Research Express, 2020, 7: 046506.

[16] LI Jia-heng, ZHANG Ying-bo, CAO Xin-yu, ZENG Qi, ZHUANG Ye, QIAN Xiao-ying, CHEN Hui. Accelerated discovery of high-strength aluminum alloys by machine learning [J]. Communications Materials, 2020, 1: 73.

[17] CHAUDRY U M, HAMAD K, ABUHMED T. Machine learning-aided design of aluminum alloys with high performance [J]. Materials Today Communications, 2021, 26: 101897.

[18] SCHMIDT J, MARQUES M R G, BOTTI S, MARQUES M A L. Recent advances and applications of machine learning in solid-state materials science [J]. NPJ Computational Materials, 2019, 5: 83.

[19] GEORGANTZIA E, GKANTOU M, KAMARIS G S. Aluminium alloys as structural material: A review of research [J]. Engineering Structures, 2021, 227: 111372. (in Chinese)

[20] BHUIYAN M S, TODA H, UESUGI K, TAKEUCHI A, WATANABE Y. Damage micromechanisms in high Mn and Zn content 7xxx aluminum alloys [J]. Materials Science and Engineering: A, 2020, 793: 139423.

[21] YUAN Ding-ling, CHEN Song-yi, CHEN Kang-hua, HUANG Lan-ping, CHANG Jiang-yu, ZHOU Liang, DING Yun-feng. Correlations among stress corrosion cracking, grain-boundary microchemistry, and Zn content in high Zn-containing Al−Zn−Mg−Cu alloys [J]. Transactions of Nonferrous Metals Society of China, 2021, 31: 2220−2231.

[22] ZOU Yan, WU Xiao-dong, TANG Song-bai, ZHU Qian-qian, SONG Hui, GUO Ming-xing, CAO Ling-fei. Investigation on microstructure and mechanical properties of Al−Zn−Mg−Cu alloys with various Zn/Mg ratios [J]. Journal of Materials Science & Technology, 2021, 85: 106−117.

[23] MUSTAFA G, GHAFFAR A, ASLAM M. A Subdivision-regularization framework for preventing over fitting of data by a model [J]. Applications & Applied Mathematics, 2013, 8: 178−190.

[24] ALEKSENDRIC D, BARTON S C, VASIC B. Prediction of brake friction materials recovery performance using artificial neural networks [J]. Tribology International, 2010, 43: 2092−2099.

[25] TAKEDA S, HAMA T, HSU H H, PIUNOVA V A, ZUBAREV D, SANDERS D P, PITERA J W, KOGOH M, HONGO T, CHENG Y. Molecular inverse-design platform for material industries [J]. Applied Data Science Track Paper, 2020, 20: 23−27.

[26] WANG Chang-sheng, FU Hua-dong, JIANG Lei, XUE De-zhen, XIE Jian-xin. A property-oriented design strategy for high performance copper alloys via machine learning [J]. NPJ Computational Materials, 2019, 5: 87.

[27] LIU Yu, ZHANG Chuan-chao, ZHANG Xie-yi, HUANG Yuan-chun. Understanding grain refinement of Sc addition in a Zr containing Al−Zn−Mg−Cu aluminum alloy from experiments and first-principles [J]. Intermetallics, 2020, 123: 106823.

[28] GUO M X, DU J Q, ZHENG C H, ZHANG J S, ZHUANG L Z. Influence of Zn contents on precipitation and corrosion of Al−Mg−Si−Cu−Zn alloys for automotive applications [J]. Journal of Alloys and Compounds, 2019, 778: 256−270.

[29] HAN Nian-mei, ZHANG Xin-ming, LIU Sheng-dan, KE Bin, XIN Xing. Effects of pre-stretching and ageing on the strength and fracture toughness of aluminum alloy 7050 [J]. Materials Science and Engineering: A, 2011, 528: 3714−3721.

[30] LI J F, BIRBILIS N, LI C X, JIA Z Q, CAI B, ZHENG Z Q. Influence of retrogression temperature and time on the mechanical properties and exfoliation corrosion behavior of aluminium alloy AA7150 [J]. Materials Characterization, 2009, 60: 1334−1341.

[31] DAI Xiao-yuan, XIA Chang-qing, LIU Chang-bin. Effect of heat treatment on microstructures and mechanical properties of Al−Zn−Mg−Cu−Sc−Zr alloys [J]. Hot Working Technology, 2005, 34: 822−825. (in Chinese)

[32] DAI Xiao-yuan, XIA Chang-qing, LIU Chang-bin. Effect of trace Sc on microstructures and properties of Al−Zn−Mg−Cu−Zr based alloys [J]. Mining and Metallurgical Engineering, 2004, 24: 59−63. (in Chinese)

[33] WANG Xiang-dong, PAN Qing-lin, LIU Li-li, XIONG Shang-wu, WANG Wei-yi, LAI Jian-ping, SUN Yuan-wei, HUANG Zhi-qi. Characterization of hot extrusion and heat treatment on mechanical properties in a spray formed ultra-high strength Al−Zn−Mg−Cu alloy [J]. Materials Characterization, 2018, 144: 131−140.

[34] FANG H C, CHAO H, CHEN K H. Effect of Zr, Er and Cr additions on microstructures and properties of Al−Zn−Mg−Cu alloys [J]. Materials Science and Engineering: A 2014, 610: 10−16.

[35] YU Zi-jian, JIN Shen-bao, FENG Man, MAXIM Y M, RUSLAN Z V, SIMON P R, SHA Gang. Temperature-dependent-composition of η phase in an Al−Zn−Mg−Cu alloy under high pressure torsion: Kinetics and thermodynamics [J]. Acta Materialia, 2022, 237: 118181.

[36] GARNER A, EUESDEN R, YAO Y C, ABOURA Y, ZHAO H, DONOHUE J, CURIONI M, GAULT B, SHANTHRAJ P, BARRETT Z, ENGEL C, BURNETT T L, PRANGNELL P B. Multiscale analysis of grain boundary microstructure in

high strength 7xxx Al alloys [J]. Acta Materialia, 2021, 202: 190−210.
[37] LIDDICOAT P V, LIAO Xiao-zhou, ZHAO Yong-hao, ZHU Yun-tian, MURASHKIN M Y, LAVERNIA E J, VALIEV R Z, RINGER S P. Nanostructural hierarchy increases the strength of aluminium alloys [J]. Nature Communications, 2010, 1: 63.
[38] ROJAS J I, SIVA B V, SAHOO K L, CRESPO D. Viscoelastic behavior of a novel aluminum metal matrix composite and comparison with pure aluminum, aluminum alloys, and a composite made of Al−Mg−Si alloy reinforced with SiC particles [J]. Journal of Alloys and Compounds, 2018, 744: 445−452.
[39] SUN Gang, WANG Shao-hua, MA Zhi-feng, LU Zheng, FENG Zhao-hui, ZHANG Xian-feng. Microstructure and mechanical properties of a new type Al−Zn−Mg−Cu alloy profile after secondary aging treatment [J]. Heat Treatment of Metals, 2012, 37: 51−54. (in Chinese)
[40] LIU S D, ZHANG X M, CHEN M A, YOU J H. Influence of aging on quench sensitivity effect of 7055 aluminum alloy [J]. Materials Characterization, 2008, 59: 53−60.
[41] CHEN K H, FANG H C, ZHANG Z, CHEN X, LIU G. Effect of of Yb, Cr and Zr additions on recrystallization and corrosion resistance of Al−Zn−Mg−Cu alloys [J]. Materials Science and Engineering: A, 2008, 497: 426−431.
[42] REN Jian, WANG Ri-chu, FENG Yan, PENG Chao-qun, CAI Zhi-yong. Microstructure evolution and mechanical properties of an ultrahigh strength Al−Zn−Mg−Cu−Zr−Sc (7055) alloy processed by modified powder hot extrusion with post aging [J]. Vacuum 2019, 161: 434−442.
[43] DAI Xiao-yuan, XIA Chang-qing, MA Ke, LIU Juan. Effect of Sc on as-cast microstructures and mechanical properties of Al−Zn−Mg−Cu−Zr alloys [J]. The Chinese Journal of Nonferrous Metals, 2007, 17: 1324−1329. (in Chinese)
[44] DITTA A, WEI L J, XU Y I, WU S J. Micro- structural characteristics and properties of spray formed Zn-rich Al−Zn−Mg−Cu alloy under various aging conditions [J]. Materials Characterization, 2020, 161: 110133.
[45] MA G N, WANG D M, LIU Z Y, XIAO B L, MA Z Y. An investigation on particle weakening in T6-treated SiC/Al−Zn−Mg−Cu composites [J]. Materials Characterization, 2019, 158: 109966.
[46] YANG S J, XING Q Y, YU J H, WANG Y L. Al−Zn−Mg−Cu alloys with strength of 800 MPa [J]. Journal of Materials Engineering, 2018, 4: 82−90.
[47] DAI Xiao-yuan, XIA Chang-qing, SUN Zhen-qi, HUA Man-yu, WU Yin. Microstructure and properties of Al−9.0Zn−2.5Mg−1.2Cu−0.12Sc−0.15Zr alloy [J]. The Chinese Journal of Nonferrous Metals, 2007, 17: 396−401. (in Chinese)
[48] HAN Nian-mei, ZHANG Xin-ming, LIU Sheng-dan, HUANG Le-yu, XIN Xing, HE Dao-guang. Influence of solution heat treatment on strength and fracture toughness of aluminum alloy 7050 [J]. Journal of Central South University (Science and Technology), 2012, 43: 855−863. (in Chinese)
[49] ZHANG Hong-tao, FU Hua-dong, HE Xing-qun, WANG Chang-sheng, JIANG Lei, CHEN Long-qing, XIE Jian-xin. Dramatically enhanced combination of ultimate tensile strength and electric conductivity of alloys via machine learning screening [J]. Acta Materialia, 2020, 200: 803−810.
[50] YAO Dan, PU Shi-long, LI Mu-yu, GUAN Ying-ping, DUAN Yong-chuan. Parameter identification method of the semi-coupled fracture model for 6061 aluminium alloy sheet based on machine learning assistance [J]. International Journal of Solids and Structures, 2022, 254: 1−13.

基于机器学习的 Al−Zn−Mg−Cu 合金快速设计

隽永飞[1]，牛国帅[2]，杨 旸[2]，徐子涵[1]，杨 健[1]，
唐文奇[1]，姜海涛[1]，韩延峰[1]，戴永兵[1]，张 佼[1,3]，孙宝德[1,3]

1. 上海交通大学 材料科学与工程学院 上海市先进高温材料及其精密成形重点实验室和金属基复合材料国家重点实验室，上海 200240；
2. 上海交通大学 计算机科学与工程系，上海 200240；
3. 上海交通大学 先进船舶和深海勘探协同创新中心，上海 200240

摘 要：提出一种基于机器学习的合金快速设计系统(ARDS)，以定制所需性能的合金制备策略或预测制备策略所对应的合金性能。为此，分别对 3 种回归算法：线性回归(LR)、支持向量回归(SVR)和人工神经网络(BPNN)进行建模和比较以训练多性能预测模型。其中，应用 SVR 构建的机器学习模型被证明是最佳的。然后，基于生成对抗网络(GAN)模型原理，构建 Al−Zn−Mg−Cu 系铝合金快速设计系统(ARDS)。对 ARDS 的预测可靠性进行验证。结果表明，为了能够获得准确的制备策略，系统中极限抗拉强度(UTS)、屈服强度(YS)和伸长率(EL)的输入上限分别约为 790 MPa、730 MPa 和 28%。此外，基于 ARDS 预测结果，制备了一种性能优异的新型铝合金材料，其 UTS 为 764 MPa、YS 为 732 MPa、EL 为 10.1%，进一步验证了 ARDS 的可靠性。

关键词：机器学习；合金快速设计系统；Al−Zn−Mg−Cu 合金；力学性能

(Edited by Wei-ping CHEN)