Artificial intelligence combined with high-throughput calculations to improve the corrosion resistance of AlMgZn alloy

Ji, Yucheng; Fu, Xiaoqian; Ding, Feng; Xu, Yongtao; He, Yang; Ao, Min; Xiao, Fulai; Chen, Dihao; Dey, Poulumi; Qin, Wentao

DOI
10.1016/j.corsci.2024.112062

Publication date
2024

Document Version
Final published version

Published in
Corrosion Science

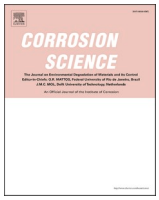
Citation (APA)
Ji, Y., Fu, X., Ding, F., Xu, Y., He, Y., Ao, M., Xiao, F., Chen, D., Dey, P., Qin, W., Xiao, K., Ren, J., Kong, D., Li, X., & Dong, C. (2024). Artificial intelligence combined with high-throughput calculations to improve the corrosion resistance of AlMgZn alloy. *Corrosion Science*, *233*, Article 112062. https://doi.org/10.1016/j.corsci.2024.112062

Corrosion Science 233 (2024) 112062

Contents lists available at ScienceDirect

Corrosion Science

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



Artificial intelligence combined with high-throughput calculations to improve the corrosion resistance of AlMgZn alloy

Yucheng Ji [a,b], Xiaoqian Fu [a], Feng Ding [a], Yongtao Xu [c], Yang He [a], Min Ao [a], Fulai Xiao [d], Dihao Chen [a], Poulumi Dey [b], Wentao Qin [a], Kui Xiao [a], Jingli Ren [e], Decheng Kong [f], Xiaogang Li [a], Chaofang Dong [a,*]

[a] Beijing Advanced Innovation Center for Materials Genome Engineering, National Materials Corrosion and Protection Data Center, Institute for Advanced Materials and Technology, University of Science and Technology Beijing, Beijing 100083, China
[b] Department of Materials Science and Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft 2628CD, the Netherlands
[c] National Engineering & Technology Research Center for Non-Ferrous Metals Composites, GRINM Group Corporation Limited, Beijing 101407, China
[d] Shandong Nanshan Aluminum Co., Ltd., Longkou 265706, China
[e] Henan Academy of Big Data, Zhengzhou University, Zhengzhou 450001, China
[f] School of Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

ARTICLE INFO

Keywords:
Al-Zn-Mg alloys
Machine learning
First-principles calculation
Molecular dynamic simulation
Precipitates

ABSTRACT

Efficiently designing lightweight alloys with combined high corrosion resistance and mechanical properties remains an enduring topic in materials engineering. Due to the inadequate accuracy of conventional stress-strain machine learning (ML) models caused by corrosion factors, a novel reinforcement self-learning ML algorithm combined with calculated features (accuracy R^2 >0.92) is developed. Based on the ML models, calculated work functions and mechanical moduli, a Computation Designed Corrosion-Resistant Al alloy is fabricated and verified. The performance (elongation reaches ~30 %) is attributed to the H trapping Al-Sc-Cu phases (-1.44 eV H^{-1}) and Cu-modified η/η' precipitates inside the grain boundaries (GBs).

1. Introduction

Al alloys, known for their lightweight properties, are widely used in aerospace [1], rail transportation [2], and electric vehicles [3] to reduce energy consumption. As the strongest 7xxx series Al alloys, its strength is attributed to the dispersion of η/η' phase (Mg-Zn) [4,5]. However, the large η/η' phase precipitates at the grain boundaries (GBs) weaken the strength of these regions [6]. When AA7xxx are utilized as structural components, they are subjected to environmental H diffusion and accumulation, resulting in stress corrosion cracking (SCC) [7]. Computations revealed that the η/η' phases at the GBs promote H aggregation, reducing the GB cohesive energy by 86.6 % [8]. Meanwhile, intergranular cracks (IGCs) are also common SCC morphologies in high-strength Al alloys [9]. Until now, the ultimate tensile strength (UTS) and elongation of SCC resistant AA7xxx reached 672 MPa and 5.02 % [10]. Even though the high strength of this alloy makes it more suitable for lightweight applications, this extremely low ductility poses a significant risk to engineering safety.

To enhance the SCC resistance of Al alloys, one approach is to modify the η/η' phases or retard their consecutive precipitation at the GBs. For instance, switching the η phases to the T phase ($Al_2Mg_3Zn_3$) by heat treatment resulted in a 60 % reduction in the cracking areal fractions [11]. Furthermore, crossover Al alloys harmonized by Mg-Zn-Cu have the potential to exhibit excellent performance in mechanical strength and corrosion, while the specific strategies are not stated [12]. High-throughput density function theory (DFT) computations provide wider chemical space for alloy design, the diffusion and stability of 86 elements in/on the Al matrix/surfaces are confirmed [13]. These calculations identified elements that can diffuse more easily than Zn or Mg thereby avoiding the formation of η/η' phase at the GBs or enhancing the cohesive energy of the GBs, such as Cu [14] and Er [15]. Nevertheless, this qualitative research only provides information about the type of beneficial elements, but the specific composition and heat treatment cannot be determined from such a study. Besides, it is also difficult to determine the reasonable range of multiple beneficial elements in a short period of time only relying on traditional methods.

* Corresponding author.
E-mail address: cfdong@ustb.edu.cn (C. Dong).

https://doi.org/10.1016/j.corsci.2024.112062
Received 8 March 2024; Received in revised form 11 April 2024; Accepted 15 April 2024
Available online 16 April 2024

Machine learning (ML) and big data may be the optimal solutions to this dilemma since they have facilitated research in the field of materials science [16]. However, the biggest problem for metal/alloy prediction is the data scale. Unlike the massive samples of image recognition or real-time speech translation [17], the alloy mechanics (corrosion) dataset size is usually limited to less than 1000 points [18,19]. To resolve this, one solution is to reuse the data or specify data labels [20]. Another option is to augment the dataset based on published work or change the predictive descriptor to more accessible features, such as hardness instead of strength [21]. In the case of high-entropy alloys, the dataset can be expanded to 5000–10,000 points using molecular dynamics (MD) simulations [22]. Exploiting calculations to extend ML datasets is a remarkable idea which can improve the prediction accuracy and generalization.

For the practically infinite chemical space for alloy design, artificial intelligence combined with physical laws brings us new opportunities [23,24]. By developing a natural language processing with deep learning, the key descriptors related to pitting potential which can be used to design corrosion-resistant alloys were confirmed [25]. Furthermore, based on random forest algorithms and phases DFT calculations, the corrosion rates of Al alloys under different environmental conditions are predicted [26]. For the highest UTS threshold, the Al fabricated with high-density η' phases reaches ~800 MPa, while its elongation is only ~5 % [27]. The performance of these phases relative to hydrogen embrittlement (HE), H trapping sites and IGCs in the Al alloys has been widely reported, and the elongation is greatly impaired in the H environment [28,29]. Thereby, optimizing the corrosion resistance of Al alloys with given phases is difficult, and there continues to be a lack of research on the global optimization of corrosion resistance and mechanical elongation. Hence, it is urgently required to propose novel strategies to design new kinds of corrosion-resistant Al alloys with high-elongation and strength.

In this study, for the mechanics Al alloys dataset with corrosion feature, ML models combined with DFT calculations are proposed to design Al alloys having attractive corrosion resistance as well as mechanical properties. The accuracy of the reinforcement learning neural network (RL-NN) model for elongation prediction achieves 0.92 (goodness of fit, R^2). Besides, the DFT-calculated formation energy, mechanical modulus, and the work function of potential phases are compared. Subsequently, the optimized composition and preferred microstructure for attractive corrosion resistance and mechanical properties are recommended based on fused ML and DFT calculations approach. Finally, a Computation Designed Corrosion-Resistant Al (CDCR-Al) alloy is manufactured to verify the reliability of the proposed design strategy. According to the slow strain rate testing (SSRT, in 0.1 M NaCl), the elongation of the CDCR-Al alloy is ~30 %. In addition, its corrosion potential is also higher than −0.7 V_{SCE}. Multi-computations reveal that the design structures are favorable for trapping H atoms, and new GB exhibits higher cracking resistance than the raw η-precipitated GB.

2. Methodology

2.1. Data standards and ML strategy

The original Al mechanical data shown in Fig. 1**a** comes from the published papers, these testing comply with the standard GB/T 228, ISO 6892 or ASTM E8 to ensure data quality. During data cleaning, data similarity and element types are checked to eliminate the same data and small sample elements (<5, e.g., Y, Hf, Nb, Cd, and La). The dataset size exceeds 1000, and the training and testing data ratios are 0.8 and 0.2, respectively.

The basic ML algorithms utilized are Back Propagation Neural Network (BP-NN) and deep learning API Keras. Fig. 1b describes the structure of strength BP-NN model, which includes the optimizer (Adam), loss function (mean square error, MSE), learning rate (0.0015), activation functions (mathematical equations that determine the output of neurons), and the neurons number. Total 20,000 epoch (forward propagation and back error correction propagation) of strength model training is performed to lower the prediction error, and layers 6 and 9 are set with a dropout parameter of 0.3 to prevent the BP-NN from overfitting. The features element and mechanical properties are linearly normalized. For the BP-NN model, the final dataset contains 34 input features, namely Zn, Mg, Cu, Zr, Ti, Si, Fe, Li, Sc, Mn, Cr, V, Ag, Ce, Eu, Er, Ni, Nd, Be, B, Sn, Pb, Sr, Na, Ca, Ga, P, Al, heat1, cool1, heat2, cool2, heat3, and corrosion (3.5 wt % NaCl). The entire heat treatment is divided into 5 components (3 for heating and 2 for cooling). For the cooling process in heat treatment (the corrosion feature is same), the feature is set to 1 if the event exists, and 0 otherwise. Learning from the Arrhenius equation, the heating process is transformed into atomic diffusion (Eq. 1).

$$heat = t\exp(T + 273)^{-1} \tag{1}$$

where T is the Kelvin temperature (K), and t indicates the holding time (10 h).

Due to the inferior corrosive elongation prediction of the conventional model, the elements combined with their contents are converted into chemical features, i.e., mass, atomic radius, electronegativity, the first ionization energy, and valence electrons. Besides, the physical features calculated by the DFT method, such as GB cohesive energy, diffusion, dissolution energy in the Al matrix, and adsorption energy on the $Al_{(100)}$ and $Al_{(111)}$, are utilized to enrich the dataset.

The index R^2 is utilized to evaluate the accuracy of ML models, which can be calculated according to Eq. 2

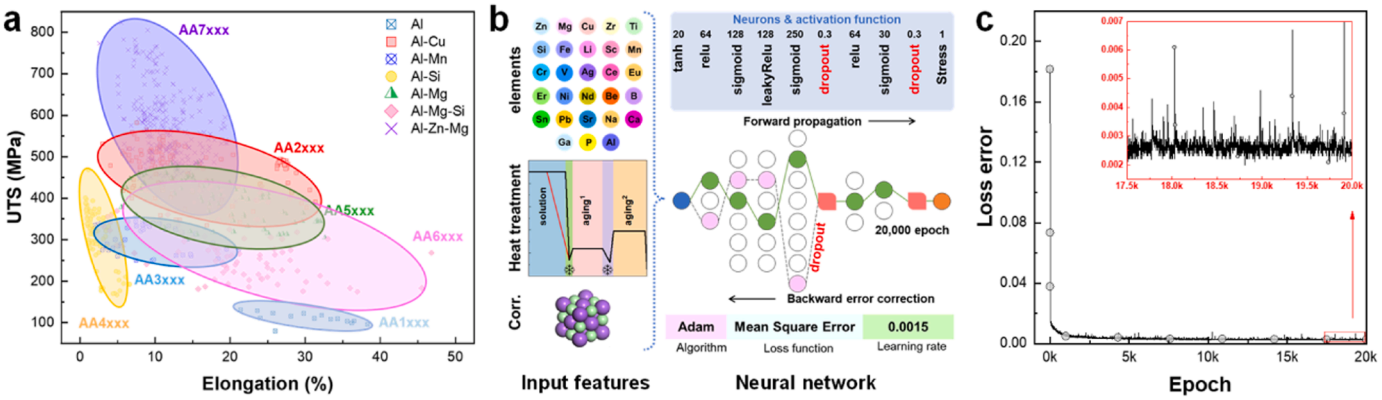


Fig. 1. Conventional ML algorithm. **a,** The raw Al alloys mechanical data categorized according to main alloying elements. **b,** The input features and loss/activation function parameters for basic BP-NN stress algorithm. **c,** The MSE loss error reduction during training.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \quad (2)$$

where $\hat{y}$, $\bar{y}_i$, and y are the predicted, averaged, and actual value of feature i.

2.2. Ab-initio calculations

All DFT calculations (9704 in total) are performed at 0 K using Vienna Ab-initio Simulation Package (VASP 5.4) [30]. The generalized gradient approximation is applied with the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [31,32]. The dissolution and surface adsorption Al models are summarized in Fig. 2, the dissolution sites include substitutional, octahedral, and tetrahedral interstitial sites. The adsorption models contain bridge, top, and HCP sites on the $Al_{(100)}$, and bridge, top, HCP and FCC on $Al_{(111)}$. The electron and force convergence accuracy are set to 10^{-5} eV and 0.01 eV $Å^{-1}$, respectively. The setting of k-points is based on the spatial length (0.2 $Å^{-1}$) in each direction, and the cutoff energy is 450 eV.

For mechanical modulus calculations, 79 kinds of phases in Al alloys are calculated according to Birch-Murnaghan equation of state (Fig. S1). The Voigt models are used to synthesize the elastic constants and generate Young's, bulk, shear modulus, and Poisson's ratio [33]. The detailed calculations are based on Eqs. 3 and 4.

$$E = WV \quad (3)$$

$$c_{ij} = \frac{\partial W}{\partial \varepsilon_i \partial \varepsilon_j} = \frac{1}{V} \frac{\partial^2 E(V, \varepsilon_k)}{\partial \varepsilon_i \partial \varepsilon_j} \quad (4)$$

where E is the DFT calculation total energy, W is the cohesive energy density, and V expresses the model volume. c_{ij} and ε_{ij} are the elastic constant and strain components, respectively.

The formation energy E_f is calculated according to Eq. 5.

$$E_f = \left(E - \sum_{i=1}^{n} E_{atom} \right) / n \quad (5)$$

where E_{atom} is the energy of single atom, and n is the total number of atoms.

The work function differences (ψ_m^p) represent the corrosion tendency of phases in Al alloys, it can be calculated by Eqs. 6 and 7.

$$\psi = \phi - E_{fermi} \quad (6)$$

$$\psi_m^p = \psi_p - \psi_m \quad (7)$$

where ϕ indicates the vacuum level determined from the average potential in the vacuum gap. E_{fermi} is the Fermi level. ψ_p (ψ_m) is the work function of the phase (pure Al matrix).

The band structures of Al_2O_3, Sc_2O_3, and Er_2O_3 are obtained using Heyd-Scuseria-Ernzerhof (HSE06) [34], which utilizes non-local exact Hartree-Fock (HF) exchange mixed with PBE to describe exchange-correlation energy (E_{XC}^{HSE}, Eq. 8).

$$E_{XC}^{HSE} = E_X^{HF,SR}(\omega) + (1-\alpha) E_X^{PBE,SR}(\omega) + E_X^{PBE,LR}(\omega) + E_C^{PBE} \quad (8)$$

where the short-range (SR) component includes the HF energy $E_X^{HF,\ SR}$ and PBE energy $E_X^{PBE,\ SR}$, while the long-range (LR) mainly is the PBE termE_C^{PBE}. α indicates the percentage of HF energy and ω is the range separation between SR and LR.

2.3. Atomic simulations

Large-scale atomic/molecular massively parallel simulator is used to perform MD and molecular statics (MS) crack modelling [35]. The MD system in Fig. 3**a** contains ~440,000 atoms, while the MS system in Fig. 3**b** is about 35,600 atoms. All systems are three-dimensional models, the size of MD and MS models are 60 × 30 × 4 nm and 28 × 25× 0.8 nm, respectively. Both GB-filled phases of MD/MS models are η-phase ($MgZn_2$, $P6_3/mmc$) and MgZnCu phase (MgZnCu, Imma). The proportions of Mg and Zn in MD models are 0.8 at % and 1.7 at %, and that of MgZnCu in MS model are both 0.9 at %. The length of pre-crack in MS model is 5 nm.

The basic ML interatomic potential method (Moment tensor potential, MTP) is developed by Shapeev groups [36], and the accuracy level of MTP is 16 with energy-weight (1), force-weight (0.01), and stress-weight (0.001). All configurations (10,122 in total) are randomly divided into training set (0.8) and validating set (0.2). Given the sufficient raw configurations, the active learning process is automatically terminated after two times (models are deforming at 800 K to generate new configurations that are supplemented to training dataset).

The isothermal-isobaric ensembles (MD models) are initialized with the temperature and pressure by velocity scaling method to 300 K and 0 MPa respectively. During deformation, the systems are maintained at 300 K, and a strain of 0.1 % is applied to the systems along the z-axis

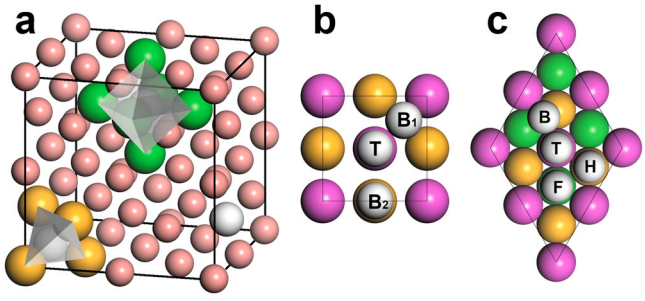


Fig. 2. The elemental DFT models that integrated in RL-NN algorithm dataset. **a,** The octahedral (green surrounded) and tetrahedral (orange surrounded) interstitial sites in Al matrix. The adsorption sites on the **b,** (100) and **c,** (111) surfaces. B indicates the bridge site, T is the top site, F indicates the FCC site, and H is the HCP site. The pink atoms are the highest layer, the orange atoms are sub-highest layer, and the green atoms are the bottom layer.

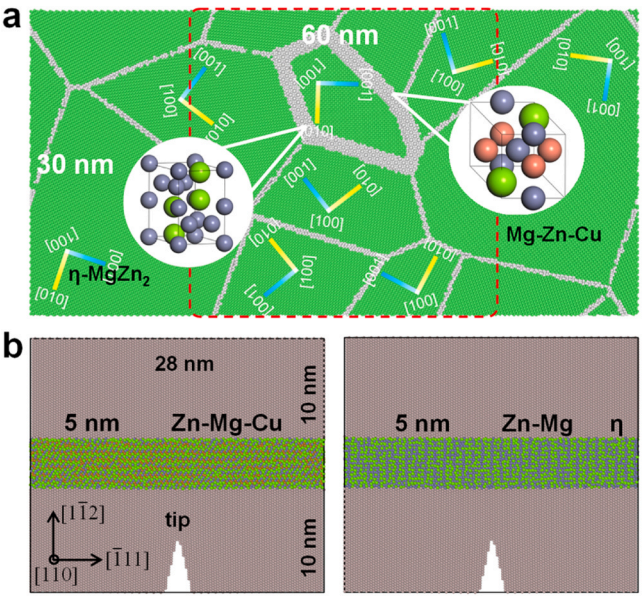


Fig. 3. Atomic models for illustrating the effects of η-phase and MgZnCu phase doped at the GB on the mechanical properties and crack propagation behavior of Al alloys: **a,** MD polycrystal model with one annular GB filled with phases. **b,** Cracking propagation MS model containing a 5 nm-thickness GB with phases.

every 1000 steps. As for MS models, the crack boundary is non-periodic and shrink-wrapped. Every step the *y*-direction (perpendicular to the cracking direction) is stretched 1 %, and then the cg and fire methods are used to minimize. The critical fracture toughness K_{Ic} and von Mises stress σ_v are calculated according to Eqs. 9 and 10.

$$K_{Ic} = 1.122\sigma\sqrt{\pi a} \tag{9}$$

where σ is the uniaxial stress and a is the crack length.

$$\sigma_v = \frac{\sqrt{2}}{2V}\left[\left(S_{xx}-S_{yy}\right)^2+\left(S_{yy}-S_{zz}\right)^2+\left(S_{zz}-S_{xx}\right)^2+6\left(S_{xy}^2+S_{yz}^2+S_{xz}^2\right)\right]^{\frac{1}{2}} \tag{10}$$

where S is the stress tensor and V indicates the atomic volume.

2.4. Verification experiments

Slow strain rate testing in corrosive environments (0.1 M NaCl and 3.5 wt % NaCl) was carried out, the strain rate was 10^{-6} s^{-1}. The thickness of SSRT sample was 3 mm, and the specific dimensions were shown in Fig. S2b. For corrosion resistance estimations, a three-electrode system (working electrode, saturated calomel electrode, and Pt counter electrode) was used to carry out the open-circuit potential, EIS, and potentiodynamic polarization measurements. The size of the working electrodes is 10 × 10 × 3 mm, and all samples were polished with 0.25 μm pastes. The EIS frequency varied from 100 kHz to 10 mHz, and the scanning rate was 10 mV min^{-1} during potentiodynamic polarization measurements. The surface Volta potential difference of nano Sc-containing phases was detected by the SKPFM (Multimode 8, Bruker), its scanning rate was 0.5 Hz with a resolution of 512×512.

Scanning electron microscopy (SEM, S-3400 N, Hitachi, Japan) combined with electron backscattered diffraction (EBSD) is used to detect the fracture morphology, grains information (size and orientations), and the propagation of secondary cracks. The EBSD samples are electron-polished in a mixed solution of 20 vol % perchloric acid and 80 vol % ethanol with a voltage of 18 V for 30 s. An FEI Talos F200X TEM is performed to observe the structure of precipitates and energy-dispersive X-ray spectroscopy. The 3DAP samples are fabricated by the focused ion beam (FEI Helios Nanolab 600i)/SEM, then they are detected in the laser mode with a pulse repetition rate of 200 kHz, and specimen temperature is 50 K.

3. Results

3.1. ML models for Al alloys with corrosion data

For the strength prediction, the conventional BBP-NN algorithms are effective [37]. By adjusting the network depth, activation functions, evaluation functions, and learning rate, a conventional Al strength BP-NN model (Fig. 1**a**) with R^2 as high as 0.967±0.007 is established within this study (Fig. 4**a**). Model accuracy is calculated based on the validation set (independent data, accounting for 20 % of the total data). For the validation of the model generalization ability, the experimental UTS and SCC strength values of AA7005, commonly utilized in high-speed trains, are 311 MPa and 284 MPa [38], respectively. The predictions of the BP-NN model for AA7005 are 326 MPa (UTS, 4.99 % error) and 287 MPa (SCC strength, 1.09 % error), respectively.

Based on the strength BP-NN model, the area over element perturbation curves (AEPC, Eq. 11) are proposed to better distinguish the elemental contribution to Al alloy strength. This method integrates the conventional mean impact value and the area over the Most Relevant First perturbation curve (MoRF, used in visual identification) [39], however, it can greatly avoid the unfair influence brought by the amount of data. The hypothesis of applying this equation is the continuous relationship between before/after perturbed UTS (see the mathematical corollary in Fig. S3). The AEPC statistical results are shown in Fig. 4**b**. Despite excluding the effect of data amount, common alloying elements (Mg, Zn, and Cu) still have considerable impacts on strength. The improvement of Zr is more obvious than that of Si and Fe, while the average area of Sc and Er reaches 2.1×10^{-3} and 1.6×10^{-4} (area index), respectively. These overall area indices greater than 10^{-4} are considered to be prominent elements in Al alloys design.

$$AEPC = \frac{1}{L}\int_{y=0}^{L}\left|f\left(x_{\pm10\%}^{k}\right)-f\left(x_{raw}^{k}\right)\right|ydy \tag{11}$$

where $f(x)$ indicates the predicted UTS via BP-NN, x represents the element content and its perturbation, k denotes the element type, y and L is the practical and the maximum strength in the Al alloy dataset, respectively.

It must be clarified that the AEPC masks the effects of elements in different strength intervals, not all attractive elements (area index $>10^{-4}$) are suitable for designing high-strength Al alloy. Hence the unintegrated results are shown in Fig. 4**c**. The 2/3 (the normalized strength) is identified as the threshold value of high-strength Al alloys, it can be seen that the Zn has the most profound impact on the UTS. Admittedly, the synergistic alloying of Zn and Mg is the most effective method to produce the high-strength Al alloys (in Fig. 1**a**, the UTS of

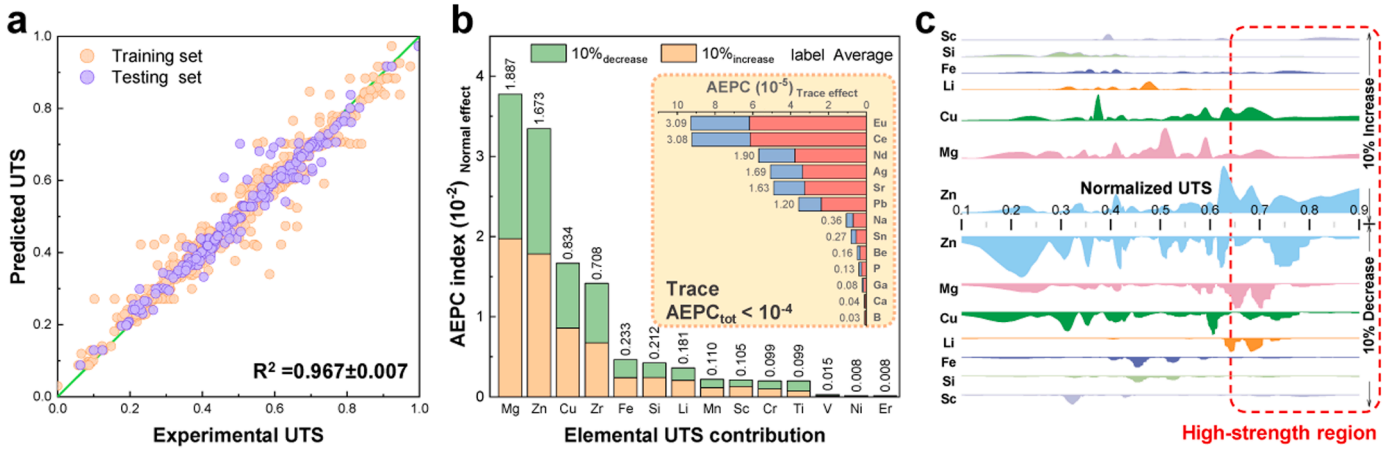


Fig. 4. Elemental strength contribution illustrated by conventional strength BP-NN model. **a,** The prediction and validation accuracies of model. **b,** AEPC elements utility ranking. **c,** the strength contribution of elements whose AEPC is larger than 0.1.

AA7xxx reaches ~800 MPa). Considering the excellent corrosion resistance of AA5xxx (Mg alloyed), the inferior corrosion resistance of AA7xxx is presumed to be related to Zn content.

The effect of Zn content on Al alloy strength is summarized in Fig. 5**d**, expectedly, the alloy strength significantly improves with Zn increases. However, the strength decreases noticeably when the Zn exceeds 12 %. There are two regions which are different from the surrounding region, namely the 2.5–3.5 % range (higher) and the 7.0–7.5 % range (lower). In the latter case (7.0–7.5 %), the mean UTS of Al alloys is only 384.3 MPa. Combined with their SCC properties (Fig. 5**b**), it is inferred that the increase of Zn content (<8 %) aggravates the SCC susceptibility of Al alloys. On the contrary, the AA7xxx with Zn content in the range of 3–4 % has better mean UTS (~413.9 MPa), and there is also no great strength degradation reported. Therefore, the Zn range of 3–4 % is a cost-effective corrosion-strength-designed Zn content range, which not only enhances the corrosion resistance of Al alloys but also reduces the density of Al alloys.

In addition, Cu, Li, Fe, and Sc also exhibit influence on the high-strength regions (Fig. 4**c**). The enrichment of Cu and Er at the GBs can improve the GB cohesive energy, which can reduce the occurrence of IGCs. Moreover, Zr likewise shows a significant effect on the strength of Al alloy, but it mainly refines grains [40] or promotes dispersed precipitation [41]. The fine grain strengthening increases the number of GBs which enlarges the risk of η/η' phases distributed at the GBs. Traditionally, Fe-containing phases are considered harmful components and are removed as much as possible during metallurgy [42], but ML shows that Fe has positive effects on the strength of Al alloy. Besides, Li is discarded owing to the DFT calculations described below.

According to elemental stability DFT calculations, Sc and Zn are overlapped in the diagram [13]. To avoid/modify the precipitation of η-$MgZn_2$ at the GBs, the diffusion barrier of elements which lower than Zn are shown in Fig. 6**a**. It can be found that diffusion ability of Sc is closed to Mg. Based on the above calculation, Sc (like Zn) and Er (improving the GB cohesive energy) are added to AA7005 to analyze their effects on mechanical and SCC properties (Fig. 6**b**). Compared with Er, the resulting improvement of Sc is more evident, particularly for UTS in the air. However, its excessive addition (>0.2 wt %) does not hugely improve their SCC performance. Despite the trace Er cannot significantly improve the UTS, its addition can improve the resistance to SCC.

Unlike the strength, the influence of the corrosive environment on

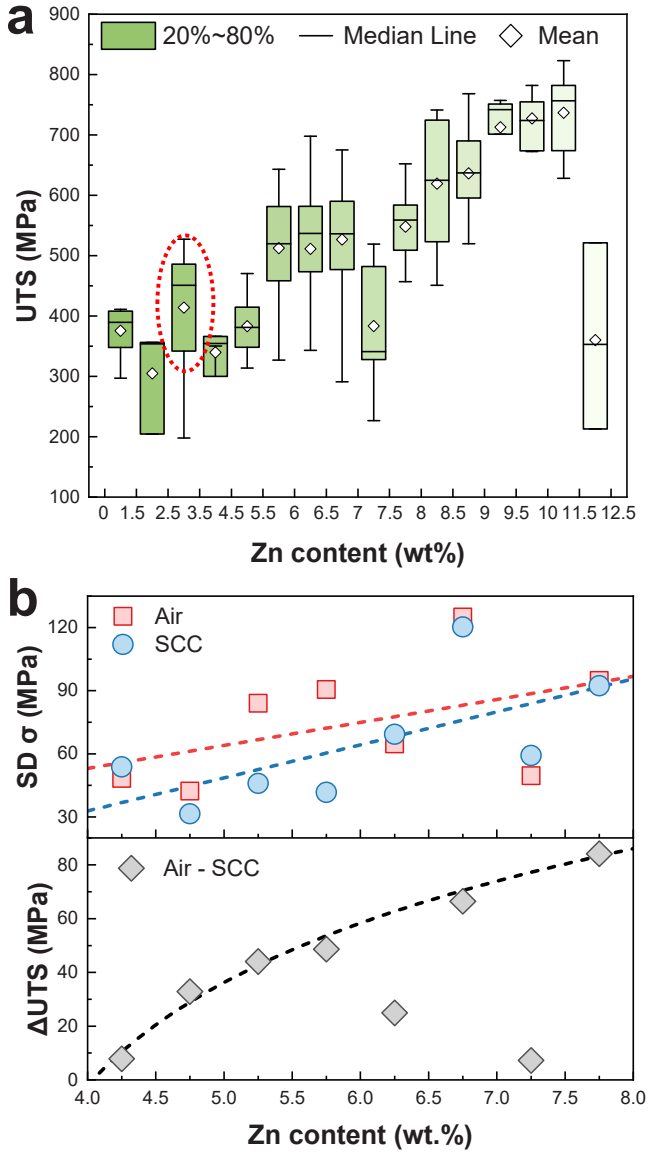


Fig. 5. Relationship between Zn content and mechanical properties of AA7xxx. **a,** Normal UTS. **b,** SCC strength susceptibility. Standard deviation σ of data and strength loss ΔUTS.

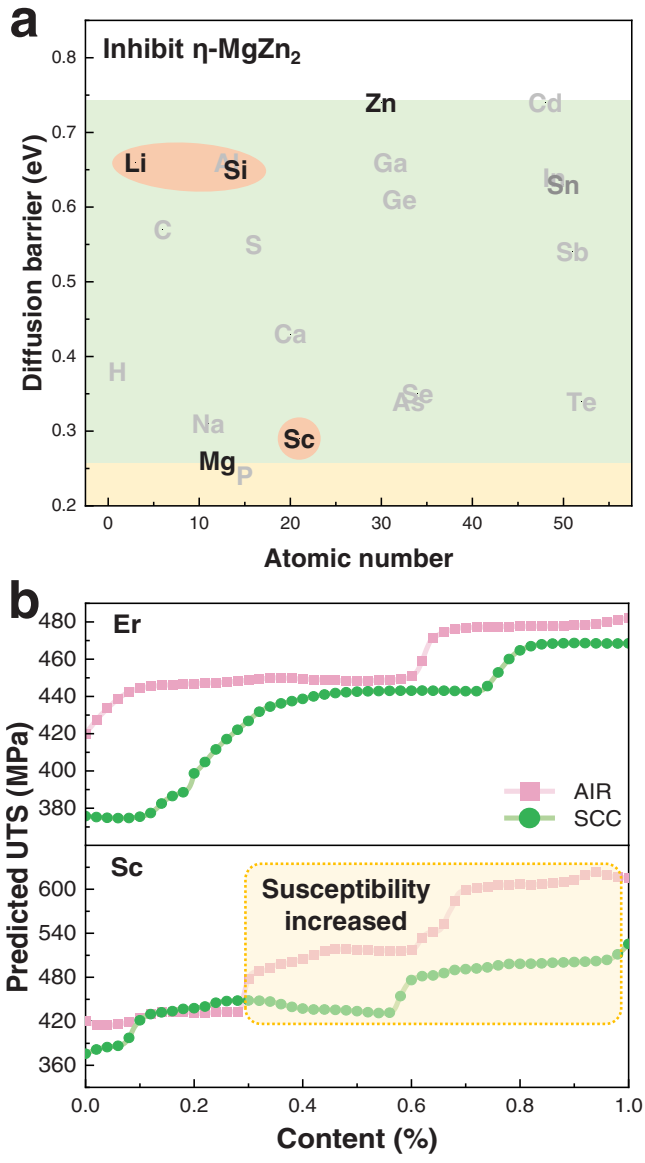


Fig. 6. Screening elements for inhibiting η-$MgZn_2$ and decreasing the SCC susceptibility. **a,** Elements whose NEB diffusion barrier in Al matrix are lower than Zn. **b,** The mechanical enhancement of Sc and Er for AA7005.

the elongation is more prominent. Quantitatively, the highest elongation prediction accuracy of the conventional BP-NN algorithm is only 0.573 ±0.09. This high error is more obvious when predicting the elongation of high-strength Al alloys, this may be attributed to the fact that corrosion promotes premature cracking and merging of cracks resulting in elongation reduction. The high accuracy of hybrid strength-elongation ML models for strength masks its disadvantage in elongation predictions, especially in corrosive environments. Hence, we propose to use DFT calculations to accurately describe the elemental physical state in the Al, and develop a reinforcement learning algorithm with them (RL-NN, Fig. 7**a**) to separately predict the elongation. The calculated features include GB cohesive energy, matrix energy barrier, chemical potential (stability), adsorbed energy on (100) and (111), matrix substitution and interstice energy (Fig. 7**b** and Fig. S4).

Except for the specialized dataset, the function pool of RL-NN model has eight activation functions and eight loss functions with the same initial accuracy probability. During each training, six activation functions and one loss function are selected according to the probability, such that a conventional neural network can be generated (the functions are randomly sorted). The training process is like the BP-NN model (Fig. 1b-c). After training, each model is evaluated based on the equations in Fig. 7**c**, the input value is the predicted elongation. Then the calculated value is fed back to the accuracy probability (accumulation). Statistically, the function probability fluctuation reveals which functions are more suitable for elongation predictions with corrosion tensile data.

As shown in Fig. 8**a**, the accuracy of the mean error functions (loss) is higher than the other types of functions. Unlike the commonly used mean square and mean absolute errors, mean squared log error (MSLE) is more suitable for the corrosion-mechanical Al alloy dataset. In terms of activation functions, the accuracy probability of the sigmoid function (commonly used in BP-NN) is exceptionally low at 9.18, ranking second to last and only higher than the softmax function. However, the activation functions, such as the linear, selu, tanh, and leakyRelu, have accuracy probabilities exceeding 10. Combining the best activation and loss functions, the model with an accuracy of 0.926±0.022 is developed (Fig. 8**b**). Its air elongation error is 8.7 %, while the error in the corrosive environment is 1.7 % (generalization ability validation). Combined with the BP/RL-NN, the integrated model has superior accuracy for the strength (0.96) and elongation (0.92) of Al alloys in air/corrosive environments.

Comparing the contribution of the traditional chemical and calculated features to the accuracy, Fig. 8**c** clearly indicates that the calculated features have higher positive influences on the RL-NN model. 5 of the top 8 features in the importance ranking are calculated features, which are GB cohesive energy, energy barrier, chemical potential,

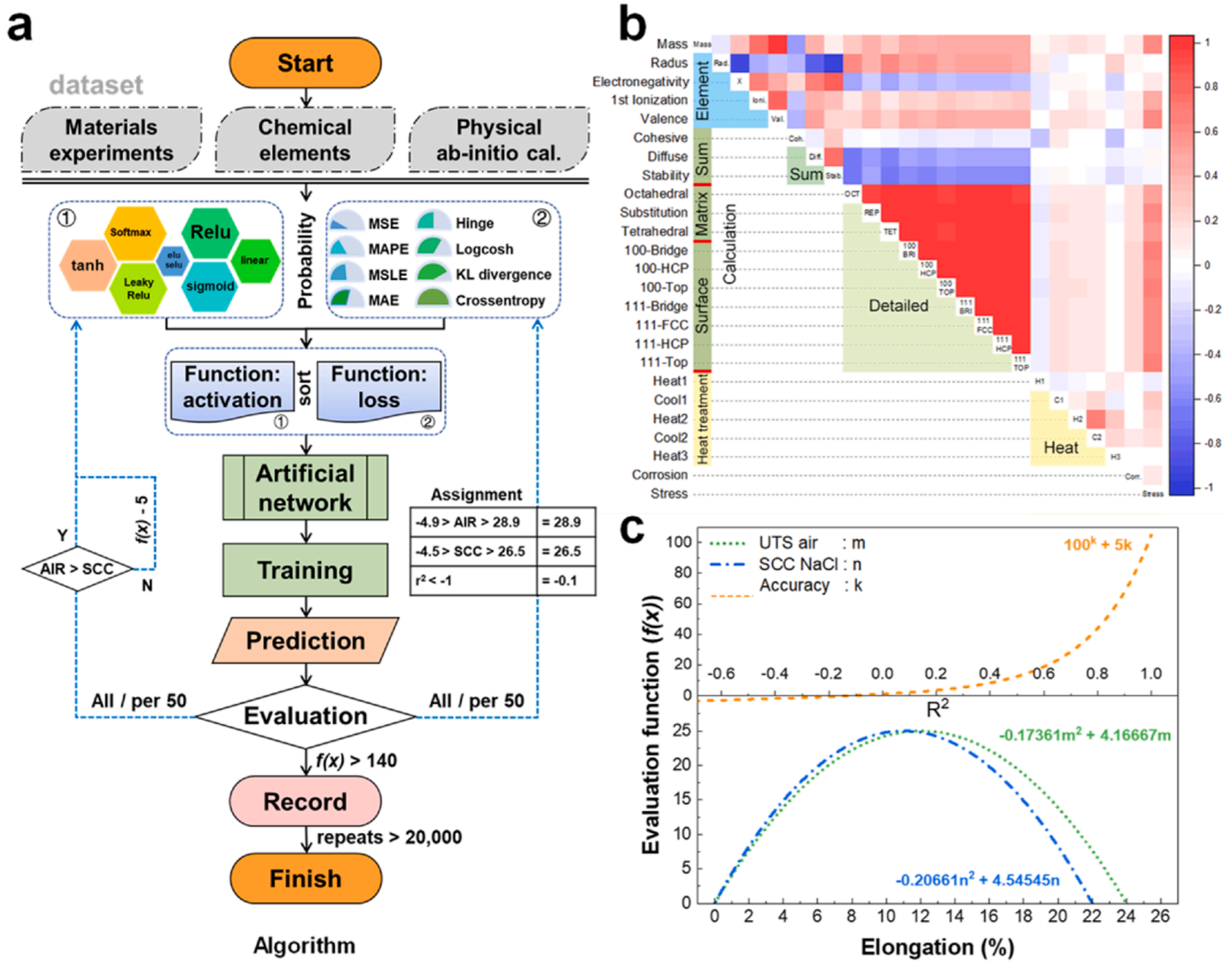


Fig. 7. Reinforcement learning algorithm for Al alloy elongation with/without corrosion. **a,** RL-NN algorithm structure. **b,** The transformed features and their Spearman correlation coefficient. **c,** Custom evaluation functions to calculate the probability of each loss/activation functions.

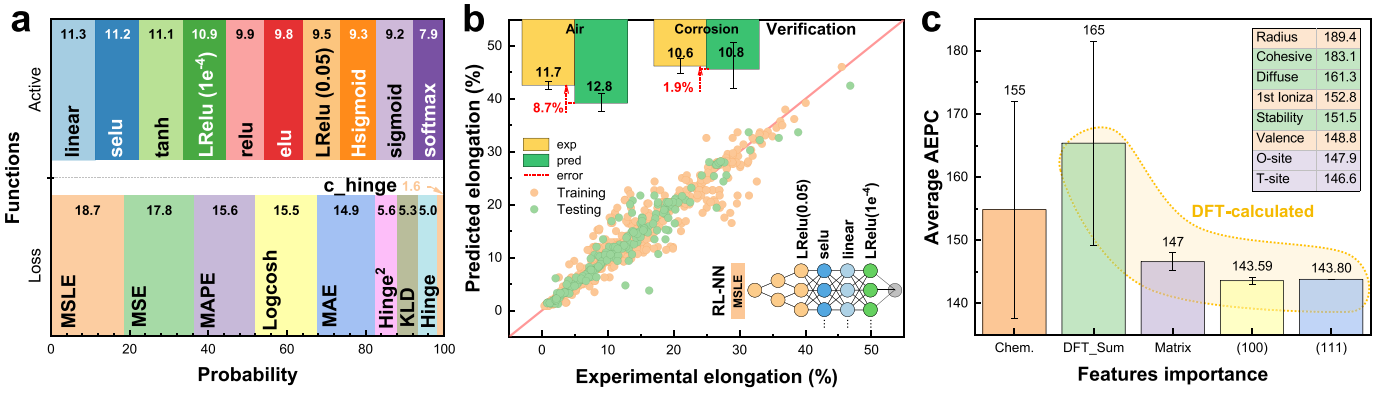


Fig. 8. Optimized RL-NN models for Al alloy elongation data. **a**, The applicability of loss and activation functions. **b,** The prediction and validation accuracies of elongation RL-NN model. **c,** The importance comparison of chemical features and DFT-calculated features.

energy in T-site and O-site. Although the most related descriptor still is chemical atomic radius which the AEPC value reaches 189.4, that of GB cohesive energy is just 3.33 % lower (183.1). Therefore, the elemental GB cohesive energies have profound impacts on the elongation prediction of Al alloys. Comprehensively, the average AEPC of the summarized calculated features (165) is greatly higher than that of the chemical features (155). After analyzing the calculated features, it also can be referred to that the calculated features related to the matrix have higher effects than the surface calculations. More specifically, the average AEPC importance of conventional calculated features is relatively low, such as that of adsorbed energy on the (100) and (111) is only 143.59 and 143.80, respectively.

3.2. DFT calculation of the secondary phases assisting in microstructural screening

While the added elements (ML determined) bring opportunities for the formation of new phases, they may also pose the risk of material degradation (mechanical properties and corrosion). Therefore, the formation energies and mechanical moduli of phases are analyzed in Fig. 9. This diagram shows calculated formation energies and modulus index (arithmetic square root of square sum of Young's modulus E and shear modulus G, $\sqrt{E^2+G^2}$, Fig. S5), and it is divided into 4 regions according to that of Al (short black dash line).

The low value of formation energy indicates that phases can form in the matrix, while an extremely low formation energy value reflects the high possibility of phase growth which may weaken the intrinsic hardening effect [43]. To determine the shape and size of the phases more precisely, the interfacial energies of phases/matrix are required. Considering the excessive computational scale of interface calculations (factors including Miller indices, misorientations, coherent, and semi/non-coherent interfaces), the formation energies are prioritized. As for the mechanical hardening effects of phases, low modulus indices cannot be directly regarded as an absence of hardening effects, since the dispersion of phases may hinder the dislocation slip.

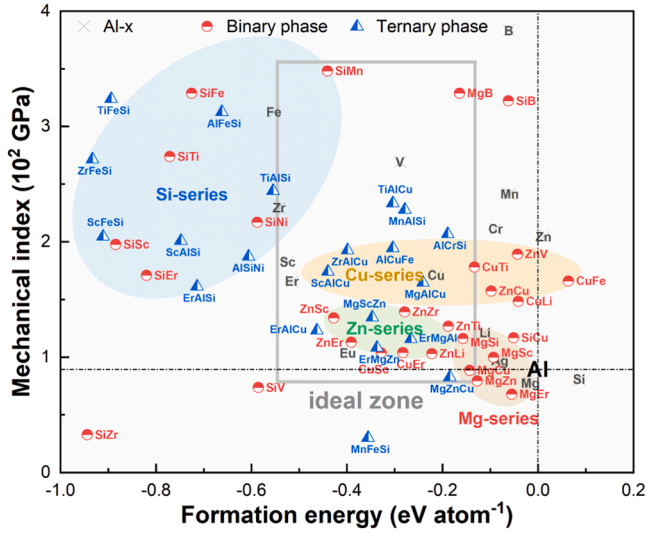


Fig. 9. Mechanical indices and formation energies of the potential secondary phases after Al alloy compositional complication.

Interestingly, the phases formed by the main alloying elements (Zn, Mg, Cu, and Si) appear to aggregate distinctly. The formation energy of the Mg group (pink) is in the range of $-0.2 \sim 0$ eV atom^{-1}, but their mechanical indices are the weakest among the main alloying elements. The Zn group (green) forms more easily than the Mg group, and their moduli indices are also greater than Mg. Pure Si will not precipitate from the matrix (≈ 0.09 eV atom^{-1}), whereas its derivatives, having lower formation energy than Zn/Mg/Cu groups, such as β [44], B′ [45], and Mg_2Si [46] can form and strengthen the alloys after heat treatment. Excessive coarsening of these Si-phases leads to a decrease in both mechanical properties [47] and corrosion resistance [48]. The Fe group exhibits the lowest formation energy among all alloying elements. Although the specific size of the phases is also governed by their interfacial energy, Fe is still unlikely to be the main alloying element due to experimentally detected harmful influence, despite its superior mechanical (Voigt statistical Young's modulus of AlFe is about 289 GPa) [49]. Therefore, the phases whose formation energies are lower than Fe (-0.55 eV atom^{-1}) may attenuate their mechanical strengthening effect and cause galvanic corrosion, therefore those phases should be avoided [50,51].

Insight into Zr-induced grains refinement (GBs proliferate and favor the distribution of η/η' phases) and Fe/Si coarsening phases, their positions are set as the left boundary of the ideal zone (gray box in Fig. 9). In addition, the accepted phases should have lower formation energy (right boundary of the gray box) and better strengthening effect (bottom boundary of the gray box) than the raw η/η' phases. To compensate for the mechanical loss (caused by the reduced Zn content), the upper left elements are more expected. Finally, the expected microstructure is to use ScAlCu, ErAlCu, or MgZnCu to replace or ameliorate the η-$MgZn_2$ phases, simultaneously, Zr, V, and B are also added to make up for mechanical strength regress. Furthermore, as shown in Table 1, the elements Er ($MgZn_2Er$, $Fm\bar{3}m$), Cu (MgZnCu, Imma), and Sc ($MgZn_2Sc$, $Fm\bar{3}m$) can significantly improve the mechanical index of raw η-$MgZn_2$ precipitates, their improvements reach 44.06 %, 47.72 %, and 71.88 %.

The corrosion behavior of the phases is another aspect to be optimized. Fig. 10 shows the work functions of various phases, and their

Table 1
Mechanical improvement of $MgZn_2$ modified by Sc, Cu, and Er.

Phase	VRH	Bulk modulus	Young's modulus	Shear modulus	Poisson's ratio	$\sqrt{E^2+G^2}$
		GPa	GPa	GPa	/	GPa
$MgZn_2$	Voigt	64.489	74.233	28.373	0.30815	79.47
	Reuss	62.989	60.098	22.408	0.34098	64.14
	Hill	63.739	67.243	25.391	0.32417	71.88
$MgZn_2Sc$	Voigt	68.067	124.22	51.94	0.19583	134.64
	Reuss	68.067	111.30	45.337	0.22747	120.18
	Hill	68.067	117.85	48.639	0.21144	127.49
MgZnCu	Voigt	44.633	76.542	31.520	0.21418	82.78
	Reuss	125.39	119.20	44.425	0.34156	127.21
	Hill	85.012	99.155	37.973	0.30561	106.18
$MgZn_2Er$	Voigt	64.033	100.39	40.520	0.23872	108.26
	Reuss	64.033	91.814	36.405	0.26102	98.77
	Hill	64.033	96.138	38.462	0.24977	103.55

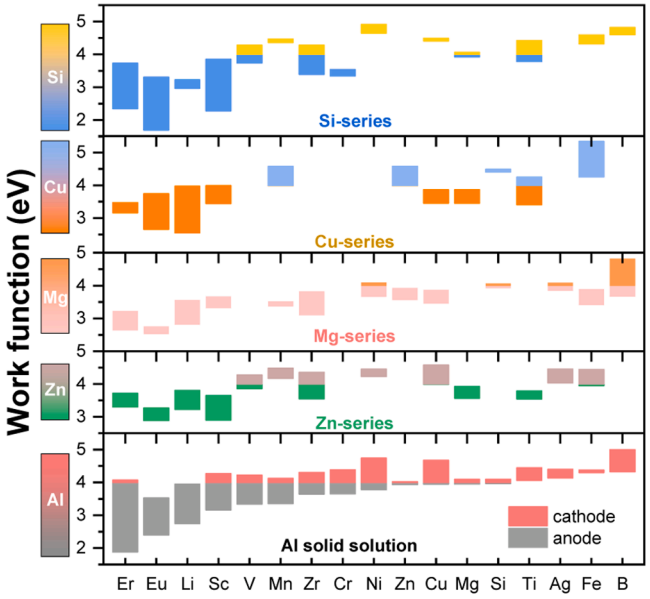


Fig. 10. The work function differences of potential secondary phases in Al alloys.

positive and negative regions are identified by Al surface (111) which exhibits higher importance than (100) in the RL-NN model. Fundamentally, an excessively low work function tends to be subjected to pitting corrosion, while a high work function risks galvanic corrosion [52,53]. From the work function in Al solid solution, it can be seen that the selected Sc or Er is not conducive to the improvement of corrosion resistance, whose lowest work functions are 3.16 eV and 1.88 eV respectively. This warns that the growth of Sc/Er-related phases must be inhibited. Besides, Fe can improve the work function of Al matrix (4.30 eV). Given the lowest formation energy of Fe group, superfluous Fe will form large-sized cathode particle resulting in galvanic corrosion [54]. As for Li, its combination with any other element generates phases with extremely low work function (2.75 eV), which are also fatal to corrosion resistance.

The work function of the Zn (2.88–4.57 eV) or Mg (2.53–4.81 eV) groups is narrow, regardless of whether the highest/lowest work function is closer to that of the Al matrix. Whereas, the phases formed by Cu (2.55–5.36 eV) or Si (1.67–4.92 eV) are difficult to be solely classified as cathode or anode, where the work function range is relatively large [51]. Meanwhile, Sc and Er approved by the aforementioned BP/RL-NN model can improve their lowest work functions by doping with Cu, which are 9.15 % and 68.37 % respectively [55]. Therefore, it can be inferred from calculations that Zn and Mg are less harmful, Fe, Sc, and Er must be precisely added to avoid the formation of the harmful phases, and Li is completely discarded. Partial alloyed Cu is expected to combine with Sc/Er to prohibit the pitting corrosion and the other Cu is desired to diffuse into the GBs.

3.3. Design and performance evaluation of Al alloy

Based on the identified element/secondary phases, the trained BP/RL-NN models and genetic algorithm are used to determine element compositions and heat treatment. The optimization goal is synergetic UTS and elongation in corrosive environments (corrosion feature is 1). Considering the high dimensionality and extremely wide composition range for the genetic algorithm (full-space search), we perform optimization for AA7005 (Table 2). The strategy is to reduce Zn and increase Cu, Fe, and Ti. Simultaneously, small amounts of Sc/Er (decrease the SCC susceptibility) and trace amounts of Zr, V, and B (compensating mechanical properties) are added. The composition restriction in Fig. 11 is divided into four gradients, namely main alloying (broadly, Zn, Mg, Cu, and Si), important but need to be strictly limited (<0.5 %, Fe), trace (<0.3 %, Sc, Mn, Ti, and Er) and sub trace (<0.1 %, Zr, V, and B).

After 100 generations, the composition of the optimal alloy (Table 2, **CDCR-Al**) converges, even in a high probability mutation environment (3 %). The determined heat treatment features in BP/RL-NN is 0.80, 1, 0.12, 1, and 0.41, respectively (Eq. 1). Given the diffusion barrier (0.57 eV) [56] and activation energy (3.2 kcal mol^{-1}) [57] of Cu in Al, the implemented solution treatment is ~505℃ for 8 h to promote the Cu diffusing into the GBs or combining with Sc. This temperature reconciles the conventional solution temperature of Al-Zn-Mg alloys (~475℃) and Al-Cu alloys (~530℃) [58,59]. Besides, the aging process is consistent with a temperature of 120℃ for 4 h. Short aging time is also beneficial to avoid excessive formation of the η/η' phases and more stable Al-Sc phase (Cu free). Subsequently, the CDCR-Al alloy is successfully manufactured via electromagnetic stirring and die casting, producing a billet with a size ϕ80×100 mm (Fig. 11).

Employing a high solid solution temperature (505℃ for 8 h), the diffusion of Cu with relatively high diffusion barrier is more active, such that the partial Al_3Sc phases are doped with Cu. SSRTs in Fig. 12**a** are used to verify the CDCR-Al alloy resistance to SCC. Clearly, the CDCR-Al alloys in a moderately corrosive environment displayed good SCC performance, with UTS reaching 443.3±38 MPa and elongation reaching 31.4±1.4 %. This exceeds the reported results (Fig. 12**a**, green region). In terms of the most pronounced elongation reduction in SCC, the elongation of the CDCR-Al alloy is almost double that of the AA7005 currently in service.

The potentiodynamic polarization curve and electrochemical impedance spectroscopy (EIS) of the CDCR-Al alloy and AA7005 are tested to evaluate their corrosion behavior. It can be seen from Fig. 12**b** that the corrosion potential (E_{corr}) of the CDCR-Al alloy is significantly improved from the AA7005 (0.095 V_{SCE}, Saturated calomel electrode).

Table 2
Chemical composition of Al alloys (wt %).

Element	Zn	Mg	Cu	Si	Sc	Fe	Er	Mn	Zr	Ti	Al
CDCR-Al	3.37	3.36	1.32	0.272	0.26	0.211	0.14	0.132	0.076	0.054	Bal.
AA7005	4.38	1.04	0.16	0.07	/	0.18	/	0.132	/	/	Bal.

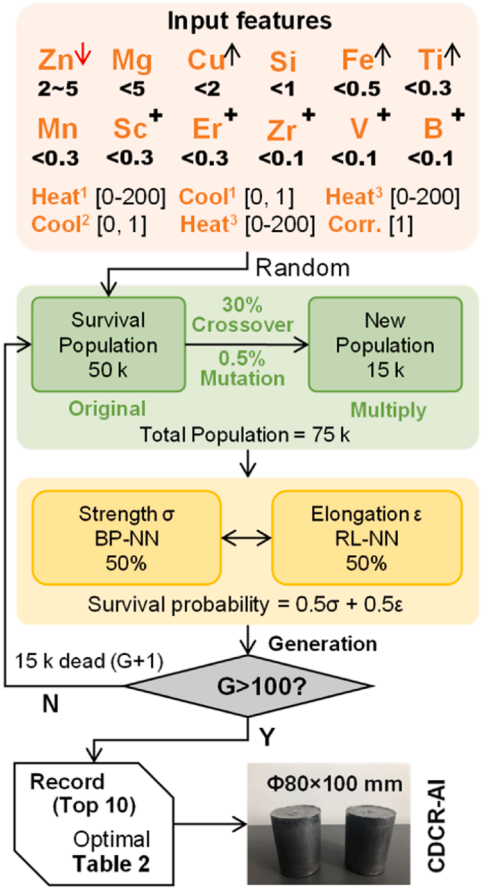


Fig. 11. Determination of CDCR-Al alloy composition and heat treatment using genetic algorithm. The value range of orange input feature are marked in black. The genetic algorithm initially generates random 50,000 alloys within the composition limitation. These alloys are multiplied and mutated to form 15,000 new alloys. Each alloy (75,000 in total) is given a survival probability by the BP/RL-NN models (predicted strength and elongation), then 15,000 alloys are eliminated according to their probabilities.

The higher E_{corr} indicates a delayed thermodynamic tendency for the anodic reaction (metal oxidation) within the same corrosion environment, while the corrosion current (i_{corr}) reflects the reaction rate. Unfortunately, not only exhibiting the lowest E_{corr}, AA7005 retains the largest i_{corr} which surges to 1.24 mA cm^{-2} when it is in the anodic stage. Analyzing the rate of O reduction reaction, it can be found from Fig. S7a that the rate of CDCR-Al alloy is slightly lower than the AA7005, the R_p of CDCR-Al alloy is 213.67 Ω cm^2, while that of AA7005 is only 18.45 Ω cm^2. This is attributed to the Volta potential difference of Cu-containing phase (cathode) with CDCR-Al alloy is 63.25 mV (Fig. S7), which is lower than the Fe-containing phase in the AA7005 (373 mV) [38]. Combined with Fig. 12**c** and **S7d**, it is found that AA7005 had an inductive reactance phenomenon at low frequency. This is attributed to the difficulty of diffusion of metal ions or other conductors. The CDCR-Al alloy exhibits two time-constants in the intermediate and low frequency. Its film impedance at the low frequency reaches 23.12 kΩ, illustrating that the surface oxide film has excellent protection (the capacitance index ~0.916). Comparing the size and depth of pitting after polarization (Fig. 12d-f), it can be found that the average pitting depth of CDCR-Al alloy is ~1.37±0.63 μm and that of AA7005 is ~8.60 ±6.03 μm. Hence, under the same environment, the corrosion resistance of CDCR-Al alloy outperforms that of AA7005.

3.4. Microstructure observation and performance explanation

From Fig. 13**a**, it is observed that there are numerous and fine phases dispersed in the grains which an average size is 36.62 nm. Combined with the energy dispersive spectroscopy (EDS) result, Al, Sc, and Cu are found. According to the formation energy shown in Fig. 9, the Al_3Sc exhibits lower formation energy than the Al-Sc-Cu phases, which indicates the Al_3Sc phases are more stable. It is found from Fig. 13**c-e** that the Al-Sc-Cu phase shows coherent relationship with the Al matrix, with $[200]_{Al}//[100]_{prep}$ and $[0\bar{1}1]_{Al}//[0\bar{1}1]_{prep}$. In addition, the interplanar spacings of phases are accurately measured as 3.86 Å and 2.77 Å, respectively. However, the length of the primitive Al_3Sc cell is confirmed as 4.105 Å [60]. When Cu replaces one Al atom in the Al_3Sc cell, the cell length can be reduced to 3.86 Å. To confirm whether the Cu doping in the Al_3Sc phases, a three-dimensional atom probe (3DAP) is performed and further confirmed that a small amount of Cu was doped into the Al_3Sc interior (Fig. 13**b**). Therefore, combined with the EDS, high-resolution transmission electron microscopy (HRTEM), and DFT calculations, we reckoned that particle Al_3Sc are doped with Cu atoms.

Furthermore, the Gibbs free energy change ΔG of H in the Al matrix and phases are computed and summarized in Fig. 13**f**. It can be clearly seen that the H captured ability of raw η/η′ phases is weakest (0.05 eV H^{-1}), which is greatly higher than that in the Al matrix. Thereby, the inner distributed η/η′ phases cannot limit the diffusion of H in grains. When the Al_3Sc phases are formed in the grains, the ΔG decreases to 0.01 eV H^{-1} which is still slightly higher than that in the matrix. However, once Cu atoms doping into the Al_3Sc, the interstice site of Al-Sc-Cu phases is greatly reduced to −1.44 eV H^{-1}. Therefore, this modification reduces the lattice parameters of phases and simultaneously enhances their H-trapping ability.

The segmented and modified GB η/η' phases (Fig. 14**b**) also significantly improve the SCC-resistance of the CDCR-Al alloy. As shown in Fig. 14**a**, the raw continuously distributed large η/η' phase at the GBs results in AA7005 significantly sensitive to the cracking, whose GB cohesive energy is only 13.4 % of raw Al GBs. However, employing the ML method, the size of the precipitation inside the GBs is decreased to 10.87±2.67 nm (Fig. 14**c**). Furthermore, the η/η' phases in the GBs are doped with Cu which are transformed into the Mg-Zn-Cu phases shown in Fig. 14**d**, Cu can enhance the GB cohesive energy (0.258 eV for Σ7 (111) GB). Based on the GB HRTEM results, it can be found from Fig. 14**e** that the precipitates continue to show a coherent relationship with the Al matrix. Consistent with the DFT calculations in Fig. 9, the Mg-Zn-Cu phase exhibits a lower formation energy than MgZn phase, yet has a higher mechanical index.

It must be reiterated, however, that not all GBs are completely transformed into new structures, such as the partial GBs in Fig. 13**a**. This may be a reason that the elongation of the CDCR-Al alloy in 3.5 wt % NaCl is 8.6 % lower than that in 0.1 M NaCl. Undesirably, Er does not apparently segregate into the GBs but dissolves into the Al matrix. Except for these, trace Er is found to combine with Sc (Fig. 14**f**). However, to detect the cracking mode and verify the effect of Mg-Zn-Cu phases on the GBs, the propagation of secondary cracks in the CDCR-Al alloy after SSRT in corrosion environment is observed by EBSD (Fig. 14**g**). Except for the primary crack, it is difficult to find secondary

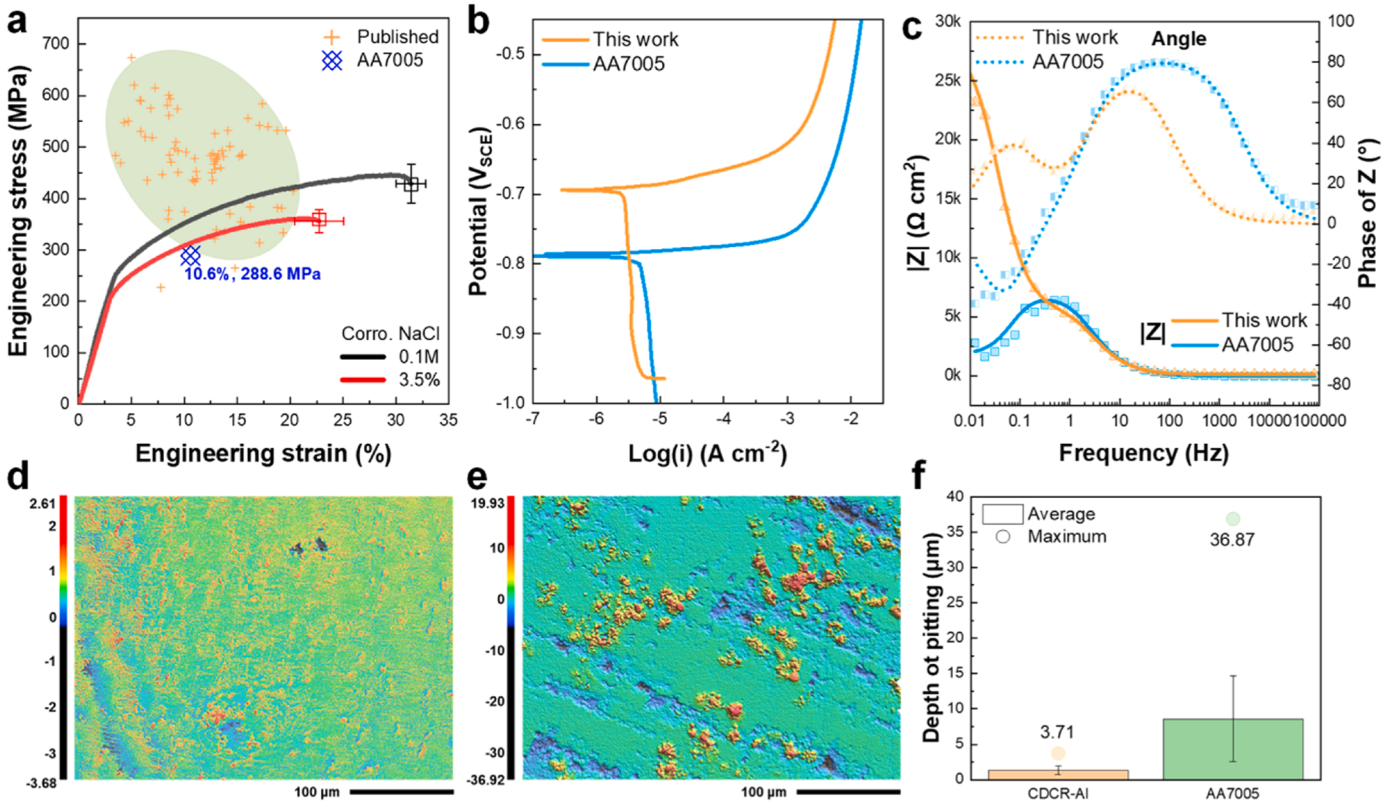


Fig. 12. Corrosion resistance evaluation of the CDCR-Al alloy. **a,** SSRTs of CDCR-Al alloy, raw AA7005 and reported works in corrosive environments. Compared with the AA7005, the stress corrosion strength and elongation of CDCR-Al alloy increase by 53 % and 193 %, respectively. **b,** Potentiodynamic polarization curve. **c,** Bode electrochemical impedance spectroscopy. The morphology of **d,** CDCR-Al alloy and **e,** AA7005 after polarization. **f,** The comparison of pitting depth. The solution for EIS testing is 0.1 M NaCl.

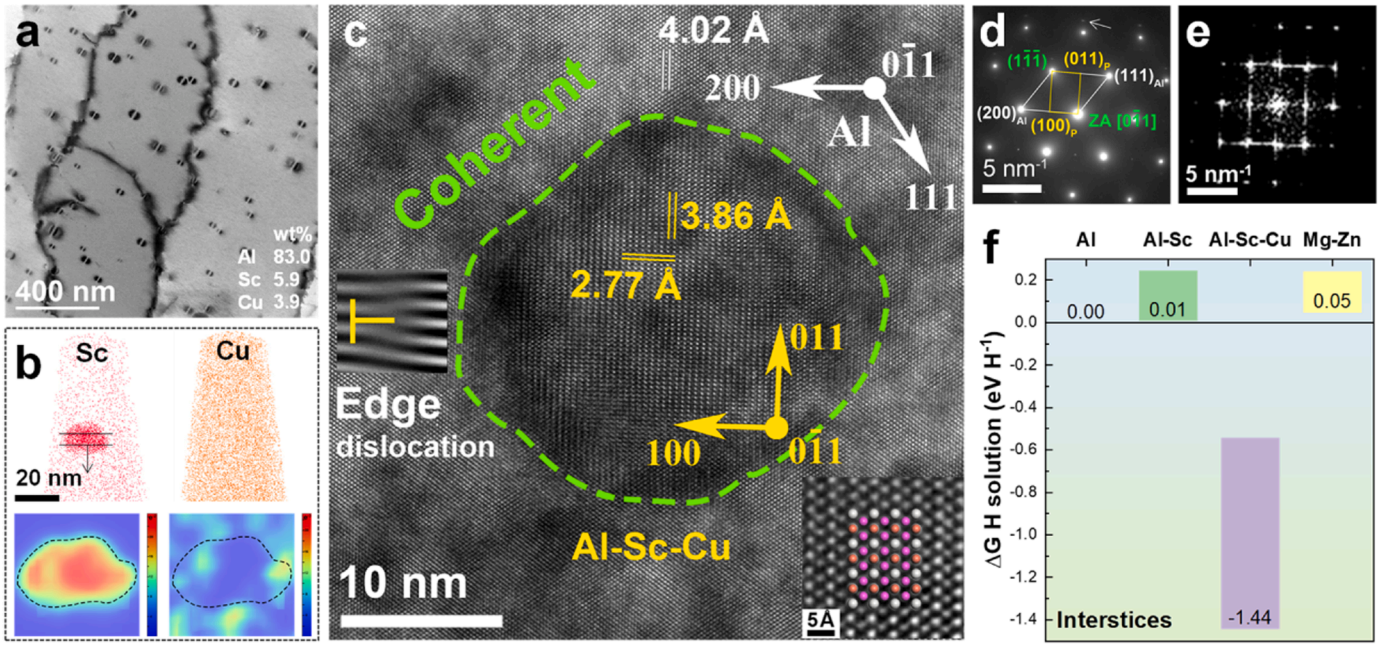


Fig. 13. Precipitates inside the grains of the CDCR-Al alloys. **a,** Dispersed fine Al-Sc-Cu precipitates. **b,** 3DAP maps of precipitates with Cu doping. **c,** Interfacial orientation relationship, **d,** diffraction pattern, and **e,** Fast Fourier Transform for precipitates. **f,** Gibbs free energy change for H trapping.

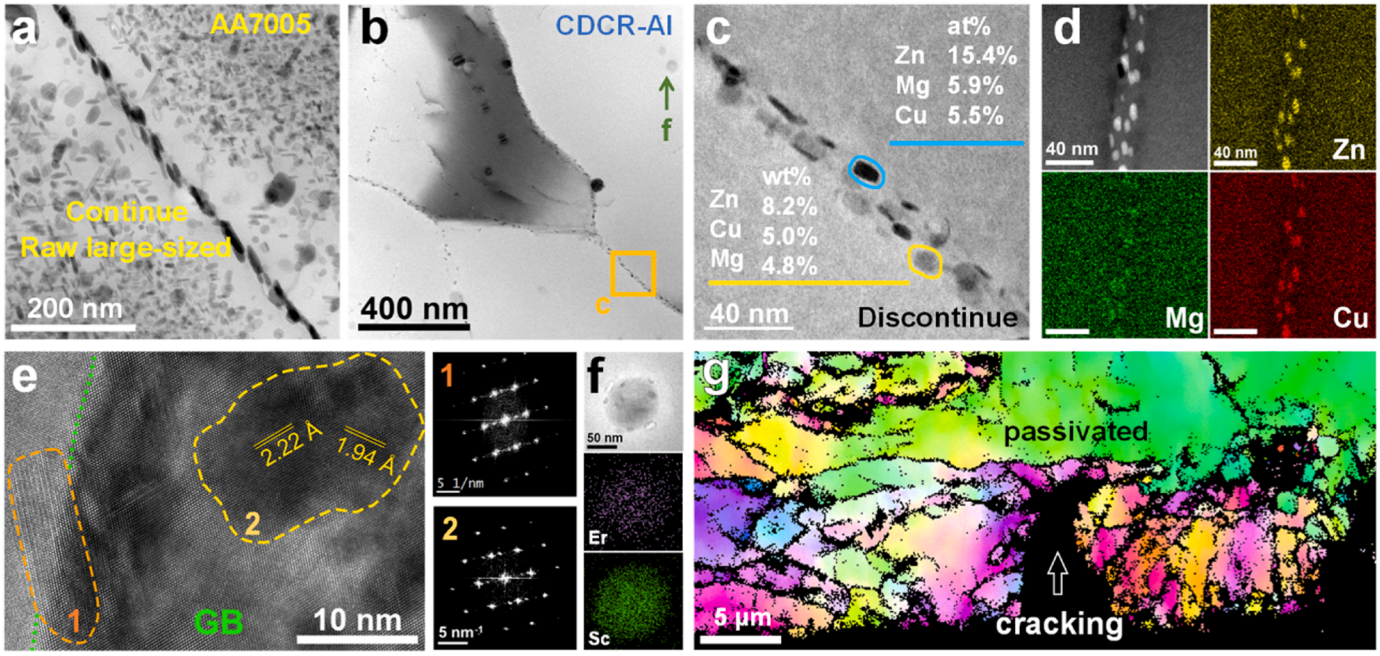


Fig. 14. Discontinue precipitated GB regions. The comparison of GBs precipitates between **a,** AA7005 and **b,** CDCR-Al alloy. **c,** Enlarged view of precipitates at the GB of CDCR-Al alloy and **d,** EDS mapping of Mg-Zn-Cu. **e,** Precipitation orientation relationship. **f,** Er is not distributed at the GBs as expected, but forms precipitates within the grains. **g,** the EBSD morphology of transgranular secondary crack without any optimization.

cracks (< 20 μm) in the alloy. Certainly, the identification rate of the sample edge is low due to the combined effect of deformation and corrosion. Combined with the SEM-detected crack position, it is found that the secondary crack in the CDCR-Al alloy propagates in a transgranular manner.

4. Discussion

Under the same corrosive environment, the corrosion current of CDCR-Al is close to that of AA5xxx, the protection of surface film is higher than that of AA7005. The average work functions of potential ternary phases are shown in Fig. 15, showing the raw Al-Sc phase exhibits extremely low work function (3.63 eV). However, when the Cu is doped to form the Al-Sc-Cu phases, the work function is greatly improved (4.18 eV), which is slightly larger than that of the Al matrix. Moreover, we utilize the SKPFM observing the height (Fig. 15**b**) and Volta potential difference (Fig. 15**c**) of nano-scale Al-Sc-Cu phase. It can be confirmed that the Volta potential difference of Cu doped Al_3Sc is extremely close to the Al matrix.

As for the complete protective film, this is attributed to the addition of Sc and Er. These two elements have stronger binding abilities to O than Al, their calculated formation energies are −3.58 eV (Sc_2O_3) and −3.71 eV (Er_2O_3), respectively. Besides, band structure calculations show that their band gaps (E_g) decrease with the enhancement of formation ability (Fig. S8). All E_g of the oxides are greater than 4.5 eV, resulting in an insulator property.

To elaborate the mechanism of Cu modified η phases to cracking, a dataset (containing 10,122 DFT configurations) is used to train an Al-Mg-Zn-Cu interatomic potential. As shown in Fig. 16**a**, its energy and force accuracies reach 6.50 meV $atom^{-1}$ and 22.19 meV $Å^{-1}$ respectively, which are lower than the report work [61]. From MD stress-strain curves at 300 K in Fig. 16**b**, it can be found that the strength of Cu-modified polycrystalline model (~5.83 GPa) is higher than η-phases (~2.21 GPa). During the tensile process, the von Mises stress of

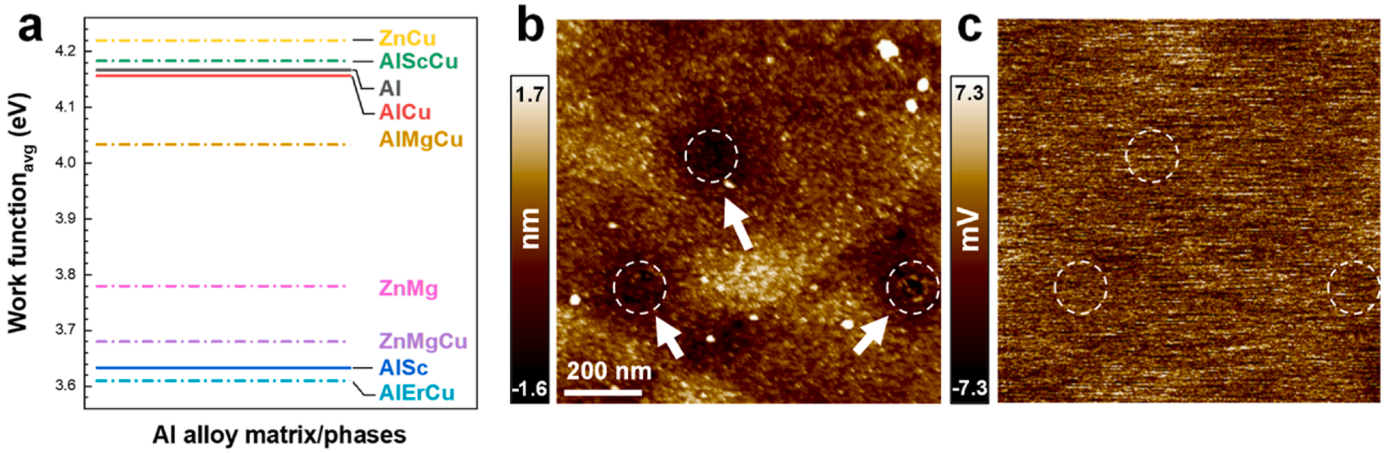


Fig. 15. Corrosion evaluation utilizing SKPFM for Al-Sc-Cu phases. **a,** average work functions of phases. The difference of Al-Sc-Cu phases in **b,** height and **c,** Volta potential.

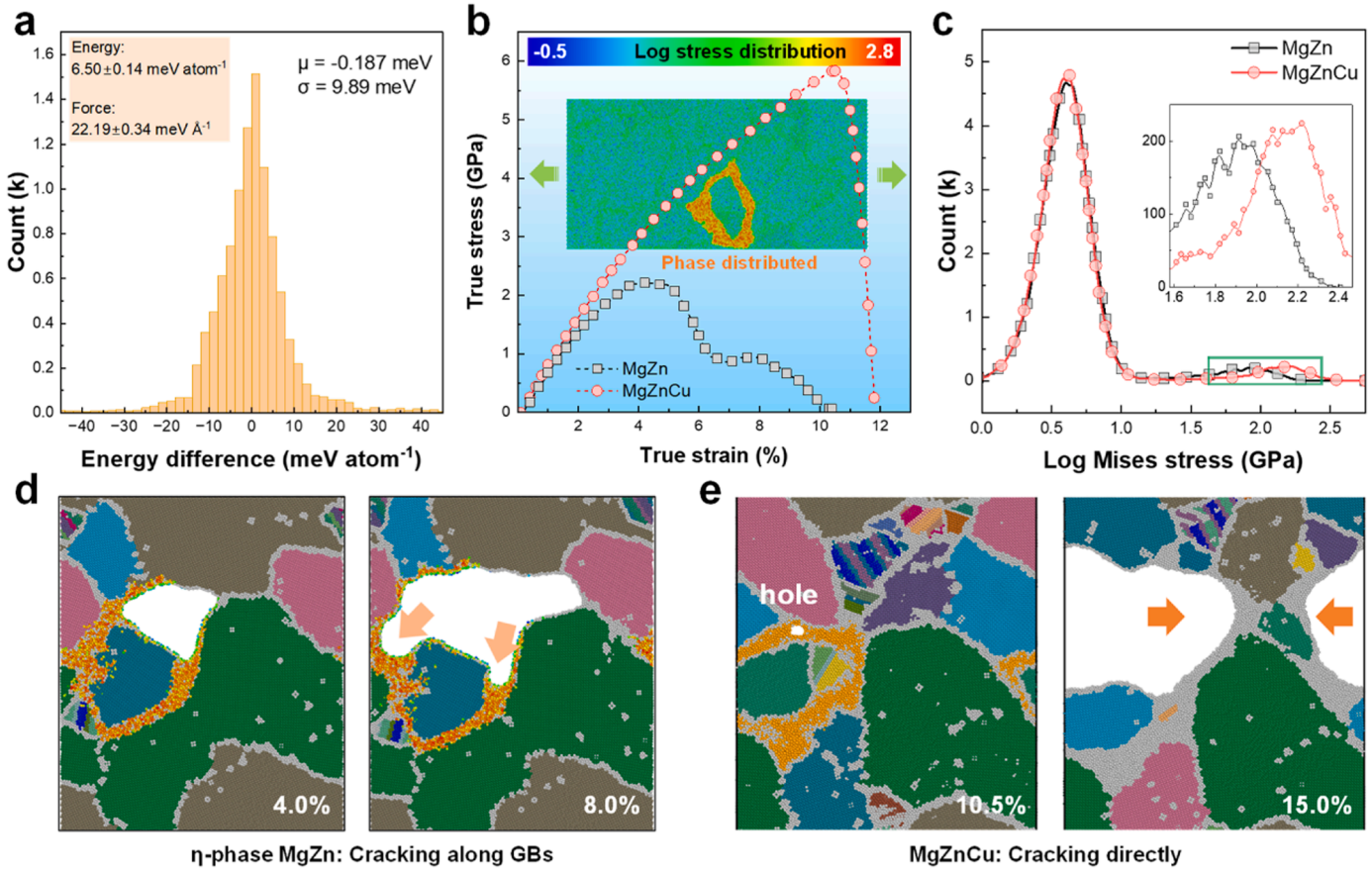


Fig. 16. MD simulations illustrating the mechanical contribution of Cu-modified η-phase on Al alloy. **a,** The energy differences predicted by Al-Mg-Zn-Cu interatomic potential and DFT. **b,** True stress-strain curves of MD tensile simulations. **c,** Von Mises stress comparison of MD models containing MgZn and MgZnCu. Failure morphologies of polycrystalline **d,** MgZn and **e,** MgZnCu models. μ is mathematical expectation of the Gaussian distribution.

precipitates at the GB (Fig. 16**c**) is greatly higher than the Al matrix (4.20 GPa). Comparing the MgZn and MgZnCu phase, the Cu modification toward η-phases further improve the von Mises stress by 101.37 %. Besides, the cracking mode of η-phases model (Fig. 16**d**) is also different from that of the Cu-modified GB model (Fig. 16**e**). When the strain reaches 4.0 %, there is a hole near the GBs, with the strain performing, it apparently propagates along the GBs at a strain of 8.0 %. While the Cu-modified model remains intact whose elongation reaches 10.5 %, and its crack propagates directly.

MS model containing single GB with a crack tip can more intuitively describe propagation of crack, therefore, the von Mises stress results are shown in Fig. 17**a-c**. It can be found that the stress at the GB is still significantly greater than the Al matrix. Eliminating the influence of temperature, we summarize the atomic von Mises stress difference according to atomic types. For conventional η-phases at the GB, it is the Mg atoms that suffer high stress which is 47.64 % higher than that of Zn atoms. When the Cu modifying the η-phase, the atomic stress distribution changes dramatically. The atom stress of Zn atoms is highest which is about 509.57 GPa, while the stress of Mg atoms decreased by 33.37 % (η-phase), to 233.51 GPa.

The strength of Cu-modified model at 0 K is like the η-phases model, but its failure elongation (63 %) is still better than that of η-phases (49 %). Owing to the high stress at the GB, the dislocations are not only emitted from the crack tip, but also occur at the GB. It can be found from Fig. 17**e** that the dislocation length of η-phases (27 nm) is lower than the MgZnCu phase (37 nm), while the number of stacking fault atoms is 59.70 % higher than that of MgZnCu phase. With the strain applied on the models, the dislocation length of MgZnCu model is significantly improved by 75 %, and the number of stacking fault atoms reach 5654. Differently, both the dislocation length and stacking fault atoms of MgZn model change slightly.

Computing the von Mises stress (Fig. 17**f**), we can find that not only the Cu-modified phases at the GB, but the stress of Al matrix (near GB) is also approximately 29.57 % greater than that of η-phases models. During the cracking, the increase rate of dislocation density of MgZnCu model is greatly larger than that of η-phases models, it indicates that the Cu-modification toward the GB hinders the slippage of dislocations. This difference also affects the critical fracture toughness K_{Ic} [62], that of MgZnCu model (0.1666 MPa $m^{-1/2}$) is ~35.56 % greater than the η-phases model, indicating the crack resistance of Cu-modified GB is stronger. Based on the SCC experiments conducted by Holroyd et al. [63], it was also confirmed that Cu addition decreased the crack growth rate by ~4 orders of magnitude. Moreover, the crack in η-phases model propagates more easily, it is fractured completely at a strain of 49 % in Fig. 17**d**, while the tip of Cu-modified GB is apparently passivated.

5. Conclusion

In this work, a strategy combining ML and DFT calculations is proposed to efficiently design an Al alloy with excellent corrosion resistance and mechanical properties. The designed structures Al-Sc-Cu nano-precipitates (in grains) and Cu modified η/η' phases (at the GBs) are beneficial to the corrosion resistance improvement of Al alloys. This investigation lead to the following conclusions:

1. Based on the strength BP-NN model ($R^2 > 0.96$), the area over element perturbation curves displays the contribution of element to strength in different intervals. Zn is advantageous for manufacturing

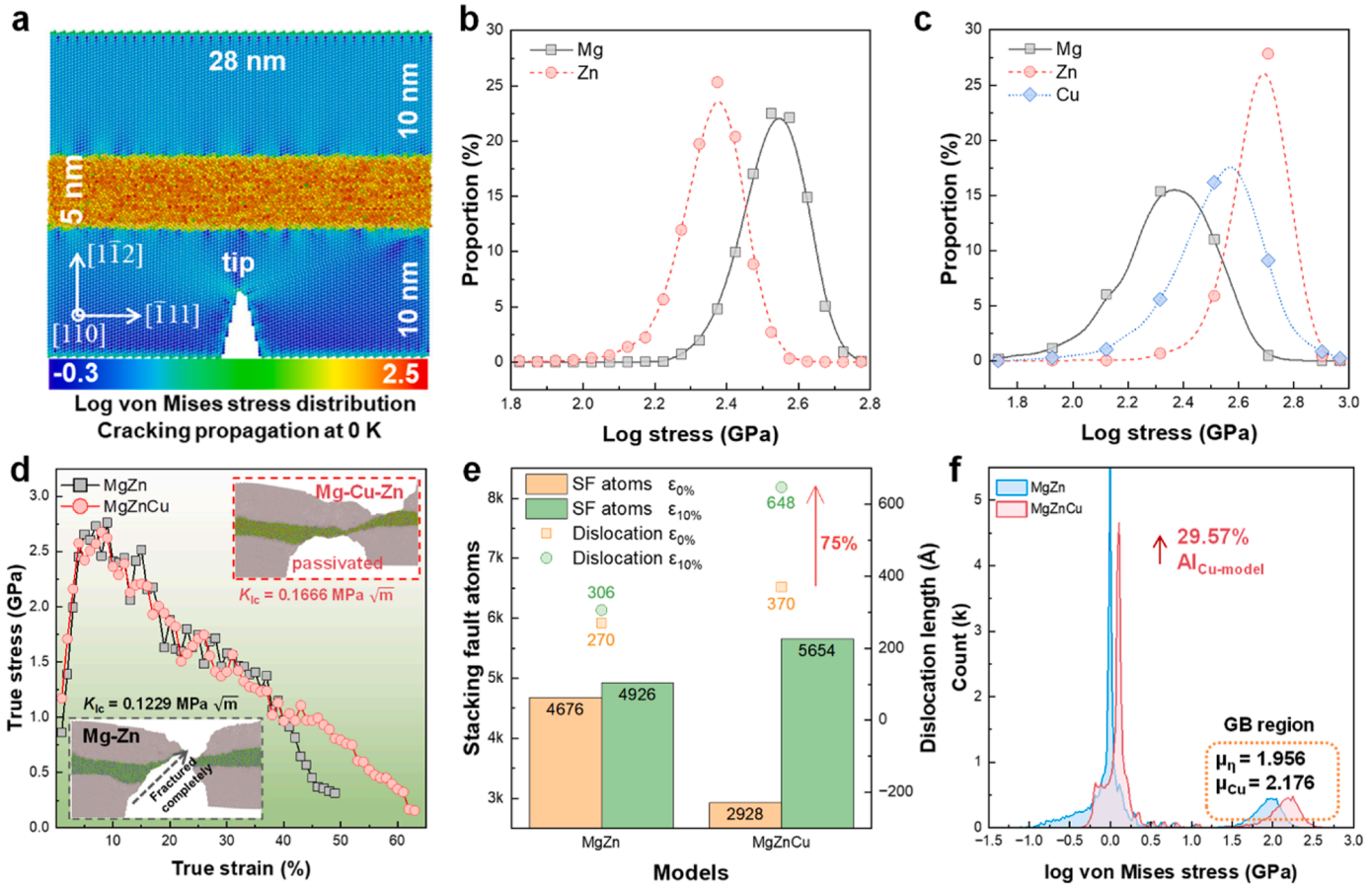


Fig. 17. MS cracking interpretation for precipitated GB of Al alloy. **a,** The von Misers stress distribution. The atomic stress statistics for **b,** MgZn and **c,** MgZnCu phases at the GB. **d,** True stress-strain curves and failure morphologies. **e,** The stacking fault atoms and dislocation length of models with strains of 0 % and 10 %. **f,** The von Mises stress comparison of MS models with 0 % strained.

high-strength Al alloys, but excessive addition makes the SCC susceptibility of Al alloys higher. Thermodynamically, NEB calculation shows that the diffusion barrier of Sc in the Al matrix is only 0.29 eV, which is close to Mg. The addition of Sc (~0.2 %) can improve the mechanical strength of Al alloys in corrosive environment.

2. Compared with the strength BP-NN model, the corrosion feature in Al alloy dataset has a greater impact on the elongation prediction. The designed reinforcement learning algorithm shows that the combination mean squared log error (loss function) and linear, selu, tanh (activation functions) is more suitable for this dataset. The DFT-calculated features can improve the accuracy of models, and they have higher average weight (~165) than normal features, especially the GB cohesive energy, energy barrier, and chemical potentials.
3. High-throughput DFT calculations screen the mechanical moduli and work function of potential secondary phases, which can quantitatively analysis the corrosion caused by phases. The computational results demonstrate the Cu-doped Al-Sc precipitates (~36.62 nm) can greatly capture the H atoms, whose ΔG for trapping H reach −1.44 eV H^{-1}. Besides, its Volta potential is close to the Al matrix. Therefore, the fabricated CDCR-Al alloy exhibits excellent mechanical properties (UTS ~450 MPa and its elongation exceeded 30 % in the 0.1 M NaCl) and corrosion potential (-0.694 V_{SCE}).
4. The modification of Cu toward η-phases at the GB can significantly improve the strength of polycrystal Al alloys, while avoid the crack propagates along the GB. Comparing the raw η-phase model, the addition of Cu decreases the atomic stress of Mg by 33.37 %, and improves the critical fracture toughness K_{Ic} by 35.56 %, its cracking tip is apparently passivated.

CRediT authorship contribution statement

Yucheng Ji: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Xiaoqian Fu:** Data curation. **Feng Ding:** Data curation. **Yongtao Xu:** Data curation. **Yang He:** Writing – review & editing, Validation. **Min Ao:** Data curation. **Fulai Xiao:** Data curation. **Dihao Chen:** Data curation. **Poulumi Dey:** Writing – review & editing. **Wentao Qin:** Investigation. **Kui Xiao:** Writing – review & editing. **Jingli Ren:** Writing – review & editing. **Decheng Kong:** Writing – review & editing. **Xiaogang Li:** Funding acquisition. **Chaofang Dong:** Writing – review & editing, Supervision, Funding acquisition, Formal analysis.

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

Data will be made available on request.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (No.52125102), Fundamental Research Funds for the Central Universities (No. FRF-TP-20-01B2), Guangdong Basic and Applied Basic Research Foundation (No. 2020B1515120093) and

Special Fund Support for Taishan Industrial Leading Talents Project. Yucheng Ji thanks for the support from China Scholarship Council (#202106460037). The authors acknowledge the platform support of the National Supercomputer Center in Tianjin and USTB MatCom of Beijing Advanced Innovation Center for Materials Genome Engineering.

Author Contributions

Y.J. compiled high-throughput calculation and machine learning algorithm, performed the experiments verification and wrote the manuscript. Y.J. and C.D. conceptualized the project and designed the research. X.F. collected and pre-processed the training data. F.D., Y.X., Y.H., M.A., and F.X. analyzed the experiment data and edited the manuscript. D.C. simulated the band structures of oxides. P.D. analyzed the DFT calculation result and edited the manuscript. W.Q. performed the SKPFM experiments. J.R. optimized the machine learning algorithm. D.K. and X.L. helped fine-tune the project and revised the manuscript.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.corsci.2024.112062.

References

[1] X. Zhang, Y. Chen, J. Hu, Recent advances in the development of aerospace materials, Prog. Aerosp. Sci. 97 (2018) 22–34.
[2] X. Sun, X. Han, C. Dong, X. Li, Applications of aluminum alloys in rail transportation, Adv. Alum. Compos. Alloy. 9 (2021) 251–268.
[3] J.C. Benedyk, Aluminum alloys for lightweight automotive structures, in: P. K. Mallick (Ed.), Materials, Design and Manufacturing for Lightweight Vehicles, Woodhead Publishing, 2010, pp. 79–113.
[4] X. Li, V. Hansen, J. Gjønnes, L. Wallenberg, HREM study and structure modeling of the η′ phase, the hardening precipitates in commercial Al-Zn-Mg alloys, Acta Mater. 47 (1999) 2651–2659.
[5] W. Sun, Y. Zhu, R. Marceau, L. Wang, Q. Zhang, X. Gao, C. Hutchinson, Precipitation strengthening of aluminum alloys by room-temperature cyclic plasticity, Science 363 (2019) 972–975.
[6] A. Garner, R. Euesden, Y. Yao, Y. Aboura, H. Zhao, J. Donoghue, M. Curioni, B. Gault, P. Shanthraj, Z. Barrett, Multiscale analysis of grain boundary microstructure in high strength 7xxx Al alloys, Acta Mater. 202 (2021) 190–210.
[7] M. López Freixes, X. Zhou, H. Zhao, H. Godin, L. Peguet, T. Warner, B. Gault, Revisiting stress-corrosion cracking and hydrogen embrittlement in 7xxx-Al alloys at the near-atomic-scale, Nat. Commun. 13 (2022) 1–9.
[8] Y. Ji, C. Dong, X. Wei, S. Wang, Z. Chen, X. Li, Discontinuous model combined with an atomic mechanism simulates the precipitated η′ phase effect in intergranular cracking of 7-series aluminum alloys, Comput. Mater. Sci. 166 (2019) 282–292.
[9] M. Lalpoor, D. Eskin, G. ten Brink, L. Katgerman, Microstructural features of intergranular brittle fracture and cold cracking in high strength aluminum alloys, Mater. Sci. Eng.: A 527 (2010) 1828–1834.
[10] L.L. Liu, Q.L. Pan, X.D. Wang, S.W. Xiong, The effects of aging treatments on mechanical property and corrosion behavior of spray formed 7055 aluminium alloy, J. Alloy. Compd. 735 (2018) 261–276.
[11] Y. Wang, B. Sharma, Y. Xu, K. Shimizu, H. Fujihara, K. Hirayama, A. Takeuchi, M. Uesugi, G. Cheng, H. Toda, Switching nanoprecipitates to resist hydrogen embrittlement in high-strength aluminum alloys, Nat. Commun. 13 (2022) 6860.
[12] L. Stemper, M.A. Tunes, R. Tosone, P.J. Uggowitzer, S. Pogatscher, On the potential of aluminum crossover alloys, Prog. Mater. Sci. 124 (2022) 100873.
[13] Y. Ji, C. Dong, L. Chen, K. Xiao, X. Li, High-throughput computing for screening the potential alloying elements of a 7xxx aluminum alloy for increasing the alloy resistance to stress corrosion cracking, Corros. Sci. 183 (2021) 109304.
[14] R.K. Gupta, A. Deschamps, M.K. Cavanaugh, S.P. Lynch, N. Birbilis, Relating the early evolution of microstructure with the electrochemical response and mechanical performance of a Cu-rich and Cu-lean 7xxx aluminum alloy, J. Electrochem. Soc. 159 (2012) C492.
[15] H. Fang, H. Chao, K. Chen, Effect of Zr, Er and Cr additions on microstructures and properties of Al–Zn–Mg–Cu alloys, Mater. Sci. Eng.: A 610 (2014) 10–16.
[16] G.L. Hart, T. Mueller, C. Toher, S. Curtarolo, Machine learning for alloys, Nat. Rev. Mater. 6 (2021) 730–755.
[17] Z. Zhou, K. Chen, X. Li, S. Zhang, Y. Wu, Y. Zhou, K. Meng, C. Sun, Q. He, W. Fan, Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays, Nat. Electron. 3 (2020) 571–578.
[18] L.B. Coelho, D. Zhang, Y. Van Ingelgem, D. Steckelmacher, A. Nowé, H. Terryn, Reviewing machine learning of corrosion prediction in a data-oriented perspective, npj Mater. Degrad. 6 (2022) 8.
[19] W. Nash, L. Zheng, N. Birbilis, Deep learning corrosion detection with confidence, npj Mater. Degrad. 6 (2022) 26.
[20] 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 (1) (2016) 9.
[21] Q. Zhao, H. Yang, J. Liu, H. Zhou, H. Wang, W. Yang, Machine learning-assisted discovery of strong and conductive Cu alloys: data mining from discarded experiments and physical features, Mater. Des. 197 (2021) 109248.
[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] X. Feng, Z. Wang, L. Jiang, F. Zhao, Z. Zhang, Simultaneous enhancement in mechanical and corrosion properties of Al-Mg-Si alloys using machine learning, J. Mater. Sci. Technol. 167 (2023) 1–13.
[24] D. Raabe, J.R. Mianroodi, J. Neugebauer, Accelerating the design of compositionally complex materials via physics-informed artificial intelligence, Nat. Comput. Sci. 3 (2023) 198–209.
[25] K.N. Sasidhar, N.H. Siboni, J.R. Mianroodi, M. Rohwerder, J. Neugebauer, D. Raabe, Enhancing corrosion-resistant alloy design through natural language processing and deep learning, Science Advances, 9 eadg7992.
[26] Y. Ji, N. Li, Z. Cheng, X. Fu, M. Ao, M. Li, X. Sun, T. Chowwanonthapunya, D. Zhang, K. Xiao, J. Ren, P. Dey, X. Li, C. Dong, Random forest incorporating ab-initio calculations for corrosion rate prediction with small sample Al alloys data, NPJ Mater. Degrad. 6 (2022) 83.
[27] J. Yong-fei, N. Guo-shuai, Y. Yang, D. Yong-bing, Z. Jiao, H. Yan-feng, S. Bao-de, Knowledge-aware design of high-strength aviation aluminum alloys via machine learning, J. Mater. Res. Technol. 24 (2023) 346–361.
[28] H. Su, H. Toda, K. Shimizu, K. Uesugi, A. Takeuchi, Y. Watanabe, Assessment of hydrogen embrittlement via image-based techniques in Al–Zn–Mg–Cu aluminum alloys, Acta Mater. 176 (2019) 96–108.
[29] H. Zhao, P. Chakraborty, D. Ponge, T. Hickel, B. Sun, C.-H. Wu, B. Gault, D. Raabe, Hydrogen trapping and embrittlement in high-strength Al alloys, Nature 602 (2022) 437–441.
[30] G. Kresse, J. Hafner, Ab initio molecular dynamics for liquid metals, Phys. Rev. B 47 (1993) 558.
[31] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (1996) 3865.
[32] B. Hammer, L.B. Hansen, J.K. Nørskov, Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals, Phys. Rev. B 59 (1999) 7413.
[33] R. Gaillac, P. Pullumbi, F.-X. Coudert, ELATE: an open-source online application for analysis and visualization of elastic tensors, J. Phys.: Condens. Matter 28 (2016) 275201.
[34] J. Heyd, G.E. Scuseria, M. Ernzerhof, Hybrid functionals based on a screened Coulomb potential, J. Chem. Phys. 118 (2003) 8207–8215.
[35] A.P. Thompson, H.M. Aktulga, R. Berger, D.S. Bolintineanu, W.M. Brown, P. S. Crozier, P.J. in't Veld, A. Kohlmeyer, S.G. Moore, T.D. Nguyen, LAMMPS-a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, Comput. Phys. Commun. 271 (2022) 108171.
[36] I.S. Novikov, K. Gubaev, E.V. Podryabinkin, A.V. Shapeev, The MLIP package: moment tensor potentials with MPI and active learning, Mach. Learn.: Sci. Technol. 2 (2021) 025002.
[37] J. Durodola, Machine learning for design, phase transformation and mechanical properties of alloys, Prog. Mater. Sci. 123 (2022) 100797.
[38] M. Ao, C. Dong, N. Li, L. Wang, Y. Ji, L. Yue, X. Sun, S. Zou, K. Xiao, X. Li, Unexpected stress corrosion cracking improvement achieved by recrystallized layer in Al-Zn-Mg alloy, J. Mater. Eng. Perform. 30 (2021) 6258–6268.
[39] W. Samek, A. Binder, G. Montavon, S. Lapuschkin, K. Müller, Evaluating the visualization of what a deep neural network has learned, IEEE Trans. Neural Netw. Learn. Syst. 28 (2017) 2660–2673.
[40] F. Wang, D. Qiu, Z.-L. Liu, J.A. Taylor, M.A. Easton, M.-X. Zhang, The grain refinement mechanism of cast aluminium by zirconium, Acta Mater. 61 (2013) 5636–5645.
[41] A. De Luca, D.C. Dunand, D.N. Seidman, Microstructure and mechanical properties of a precipitation-strengthened Al-Zr-Sc-Er-Si alloy with a very small Sc content, Acta Mater. 144 (2018) 80–91.
[42] Y. Zhu, K. Sun, G. Frankel, Intermetallic phases in aluminum alloys and their roles in localized corrosion, J. Electrochem. Soc. 165 (2018) C807.
[43] S. Cheng, Y. Zhao, Y. Zhu, E. Ma, Optimizing the strength and ductility of fine structured 2024 Al alloy by nano-precipitation, Acta Mater. 55 (2007) 5822–5832.
[44] H. Li, Z. Yan, L. Cao, Bake hardening behavior and precipitation kinetic of a novel Al-Mg-Si-Cu aluminum alloy for lightweight automotive body, Mater. Sci. Eng.: A 728 (2018) 88–94.
[45] H. Chen, J. Lu, Y. Kong, K. Li, T. Yang, A. Meingast, M. Yang, Q. Lu, Y. Du, Atomic scale investigation of the crystal structure and interfaces of the B′ precipitate in Al-Mg-Si alloys, Acta Mater. 185 (2020) 193–203.
[46] L. Lu, M. Lai, M. Hoe, Formaton of nanocrystalline Mg2Si and Mg2Si dispersion strengthened Mg-Al alloy by mechanical alloying, Nanostruct. Mater. 10 (1998) 551–563.
[47] M.W. Zandbergen, Q. Xu, A. Cerezo, G.D.W. Smith, Study of precipitation in Al–Mg–Si alloys by Atom Probe Tomography I. Microstructural changes as a function of ageing temperature, Acta Mater. 101 (2015) 136–148.
[48] S.K. Kairy, N. Birbilis, Clarifying the role of Mg2Si and Si in localized corrosion of aluminum alloys by quasi in situ transmission electron microscopy, Corrosion 76 (2020) 464–475.
[49] P. Ashtari, H. Tezuka, T. Sato, Modification of Fe-containing intermetallic compounds by K addition to Fe-rich AA319 aluminum alloys, Scr. Mater. 53 (2005) 937–942.

[50] N. Li, C. Dong, C. Man, X. Li, D. Kong, Y. Ji, M. Ao, J. Cao, L. Yue, X. Liu, Insight into the localized strain effect on micro-galvanic corrosion behavior in AA7075-T6 aluminum alloy, Corros. Sci. 180 (2021) 109174.
[51] Y. Ji, C. Dong, D. Kong, X. Li, Design materials based on simulation results of silicon induced segregation at AlSi10Mg interface fabricated by selective laser melting, J. Mater. Sci. Technol. 46 (2020) 145–155.
[52] M. Liu, Y. Jin, C. Leygraf, J. Pan, A DFT-Study of Cl Ingress into α-Al2O3 (0001) and Al (111) and its possible influence on localized corrosion of Al, J. Electrochem. Soc. 166 (2019) C3124.
[53] S. Li, G.S. Frankel, C.D. Taylor, Theoretical relations between electronic and ionic work functions, standard reduction potentials for metal dissolution and the corrosion potential, J. Electrochem. Soc. 169 (2022) 081506.
[54] D.S. Kharitonov, I. Dobryden, B. Sefer, J. Ryl, A. Wrzesińska, I.V. Makarova, I. Bobowska, I.I. Kurilo, P.M. Claesson, Surface and corrosion properties of AA6063-T5 aluminum alloy in molybdate-containing sodium chloride solutions, Corros. Sci. 171 (2020) 108658.
[55] S.K. Kairy, B. Rouxel, J. Dumbre, J. Lamb, T.J. Langan, T. Dorin, N. Birbilis, Simultaneous improvement in corrosion resistance and hardness of a model 2xxx series Al-Cu alloy with the microstructural variation caused by Sc and Zr additions, Corros. Sci. 158 (2019) 108095.
[56] M. Mantina, Y. Wang, L.Q. Chen, Z.K. Liu, C. Wolverton, First principles impurity diffusion coefficients, Acta Mater. 57 (2009) 4102–4108.
[57] N. Peterson, S. Rothman, Impurity diffusion in aluminum, Phys. Rev. B 1 (1970) 3264.
[58] L. Jiang, B. Rouxel, T. Langan, T. Dorin, Coupled segregation mechanisms of Sc, Zr and Mn at θ′ interfaces enhances the strength and thermal stability of Al-Cu alloys, Acta Mater. 206 (2021) 116634.
[59] R.F. Gutiérrez, F. Sket, E. Maire, F. Wilde, E. Boller, G. Requena, Effect of solution heat treatment on microstructure and damage accumulation in cast Al-Cu alloys, J. Alloy. Compd. 697 (2017) 341–352.
[60] A.F. Norman, P.B. Prangnell, R.S. McEwen, The solidification behaviour of dilute aluminium–scandium alloys, Acta Mater. 46 (1998) 5715–5732.
[61] D. Marchand, W. Curtin, Machine learning for metallurgy IV: a neural network potential for Al-Cu-Mg and Al-Cu-Mg-Zn, Phys. Rev. Mater. 6 (2022) 053803.
[62] Y. Sun, Q. Peng, G. Lu, Quantum mechanical modeling of hydrogen assisted cracking in aluminum, Phys. Rev. B 88 (2013) 104109.
[63] N.H. Holroyd, G. Scamans, Stress corrosion cracking in Al-Zn-Mg-Cu aluminum alloys in saline environments, Metall. Mater. Trans. A 44 (2013) 1230–1253.