

Research Article

Open Access



# Multi-model integration accelerates Al-Zn-Mg-Cu alloy screening

Yanru Yuan<sup>1,2</sup>, Yudong Sui<sup>1,2,\*</sup> , Pengfei Li<sup>1,2</sup>, Meng Quan<sup>1,2</sup>, Hao Zhou<sup>3</sup>, Aoyang Jiang<sup>4,\*</sup>

<sup>1</sup>Faculty of Materials Science and Engineering, Kunming University of Science and Technology, Kunming 650093, Yunnan, China.

<sup>2</sup>National & Local Joint Engineering Laboratory of Advanced Metal Solidification Forming and Equipment Technology, Kunming University of Science and Technology, Kunming 650093, Yunnan, China.

<sup>3</sup>Institute of Materials Plainification, Liaoning Academy of Materials, Shenyang 110167, Liaoning, China.

<sup>4</sup>School of Electrical Engineering, Sichuan University, Chengdu 610000, Sichuan, China.

\***Correspondence to:** Prof. Yudong Sui, Faculty of Materials Science and Engineering, Kunming University of Science and Technology, Kunming 650093, Yunnan, China. E-mail: [suiyd2016@126.com](mailto:suiyd2016@126.com); Mr. Aoyang Jiang, School of Electrical Engineering, Sichuan University, Chengdu 610000, Sichuan, China. E-mail: [1074283118@qq.com](mailto:1074283118@qq.com)

**How to cite this article:** Yuan Y, Sui Y, Li P, Quan M, Zhou H, Jiang A. Multi-model integration accelerates Al-Zn-Mg-Cu alloy screening. *J Mater Inf* 2024;4:23. <https://dx.doi.org/10.20517/jmi.2024.34>

**Received:** 9 Aug 2024 **First Decision:** 14 Sep 2024 **Revised:** 4 Nov 2024 **Accepted:** 9 Nov 2024 **Published:** 19 Nov 2024

**Academic Editors:** Qian Li, Hao Li **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

## Abstract

The 7xxx (Al-Zn-Mg-Cu) alloys were extensively utilised in aerospace and rail transit due to their low density, high strength, and excellent processability. Nevertheless, an increase in strength inevitably comes at the expense of ductility, and vice versa, which was known as the “strength-ductility trade-off” dilemma. Extensive research had been conducted to address this issue, accumulating a large amount of experimental and computational data. However, the conventional approach of trial-and-error largely relied on the empirical knowledge of researchers. If the expected results are not obtained, the experiment will need to be repeated constantly. The emergence of machine learning (ML) as a new paradigm had introduced novel data analysis tools for basic scientific research. Numerous studies had demonstrated the feasibility and reliability of ML methods. Consequently, this paper adopted ML methods to investigate this dilemma of Al-Zn-Mg-Cu alloys. A multi-algorithm integrated model was employed to accelerate the screening of the target mechanical properties. Four high-strength and high-ductility alloys were designed using the desired target properties as inputs. The alloys designed by the ML methods were experimentally prepared. Additionally, the effects of extrusion and rolling processes on the alloy’s properties were compared. Notably, the E4 alloy exhibited an ultimate tensile strength (UTS) of  $709 \pm 4$  MPa and an elongation (EL) of  $16\% \pm 1\%$ , which represents a significant enhancement in comprehensive performance. This study provides



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



a reference for resolving the “strength-ductility trade-off” dilemma, and contributes to the development of more competitive aluminium alloy materials.

**Keywords:** Al-Zn-Mg-Cu alloy, machine learning, ensemble learning, mechanical properties

## INTRODUCTION

Aerospace manufacturing is a critical reflection of the national industrial manufacturing level. For a long time, aluminium alloys have been widely used in structural components such as spacecraft rocket fuel tanks, airframes, and large sealed capsules, which are indispensable key structural materials in the modern aerospace field<sup>[1-3]</sup>. Currently, high-performance materials are being iterated as the aerospace industry advances, and the emergence of new materials is driving traditional materials forward. Aluminium alloys are facing great challenges and are focusing on balancing the properties of many aspects to meet the comprehensive requirements of the future aerospace field<sup>[4]</sup>. The 7xxx aluminium alloy is a typical aerospace aluminium alloy, which is widely used because of its ultra-high tensile strength often exceeding 500 MPa<sup>[5]</sup>, such as 7075, 7475, 7050, *etc.* Nevertheless, its low ductility makes it easily cracked and deformed during the forming and processing restricting the application. However, strength and ductility are usually mutually exclusive in aluminium alloys. In other words, the increase of strength will inevitably come at the expense of ductility and vice versa. This is the so-called “strength-ductility trade-off” dilemma<sup>[6,7]</sup>.

Researchers usually seek a balance between strength and ductility by optimising alloy compositions, heat treatment processes, forming processes and other methods. In particular, a great deal of research has been conducted on alloy compositions. It is widely acknowledged that the main components Zn, Mg, and Cu play a decisive role in the alloy’s microstructure and properties. Li *et al.* demonstrated that by simultaneously altering the Cu content and the Zn/Mg ratio, the tensile strength rose by 5.9% with a 0.21% increase in Cu content, while the elongation (EL) improved by 8.7 % with a 0.59% increase in Zn/Mg ratio<sup>[8]</sup>. Zhang *et al.* investigated the impact of Cu concentration on the mechanical properties of alloys<sup>[9]</sup>. The developed Al-4.0Mg-3.0Zn-1.5Cu alloy illustrates an excellent combination of strength and ductility. Its yield strength (YS) increases by 80.5% without loss of ductility as compared to the Cu-free alloy. By regulating the alloy composition and introducing selectively microalloying elements such as Ni, Zr, Mn, Cr, Sc, Ce, Y, *etc.*<sup>[10-17]</sup>, the microstructure of the alloy can be significantly altered, leading to a reasonable distribution of phases that effectively improves the comprehensive mechanical properties. Studies<sup>[18-24]</sup> have shown that these elements tend to precipitate preferentially as nucleation points for grain growth during the solidification of alloys facilitating the refinement of grains. Furthermore, these elements also form dispersed secondary phases, which impede the movement of dislocations at the microscopic level enhancing the material properties further.

After long-term research, a large amount of experimental and computational data has been accumulated, which is massive and redundant. However, traditional metal material design mainly relies on a large number of experimental experiences, assisted by guiding computational simulation. If the desired effect is not achieved, it is necessary to constantly adjust and experiment through trial and error. The research cost is high and the cycle is long, which is difficult to meet the high efficiency and low-cost design of high-end key metal materials. It takes about 10-20 years to reach the final industrial application, and the iteration of new materials lags far behind the product design. With the arrival of the big data era, the emergence of machine learning (ML) as the representative of the scientific research fourth paradigm can mine the laws in big data effectively based on the relevant algorithms and do not need to understand the inherent mechanism, which has become a new starting point for industrial design. Numerous studies<sup>[25-29]</sup> have demonstrated the feasibility and reliability of ML methods for designing alloys with excellent properties. Chaudry *et al.* used



ML models to identify aluminium alloys with excellent hardness<sup>[30]</sup>. Predicting alloy hardness is based on composition and process parameters, *etc.* The results showed that the Gradient Boosted Tree model performed the best with a coefficient of determination ( $R^2$ ) of 0.94. Wang *et al.* proposed a ML design system (MLDS) that learns the relationship between the strength and conductivity of materials and alloy compositions in order to screen the alloy compositions<sup>[31]</sup>. Eventually, two sets of alloys were prepared based on the ML reverse design results with an error of less than 6% between the measured and target properties. Chen *et al.* introduced a ML-assisted strategy that simultaneously enhancing the strength and ductility of the alloy by 27% and 13.5%, respectively, by optimising the two-step aging treatment parameters with only four new sequential experiments<sup>[32]</sup>. Mi *et al.* created a reverse ML design model and designed four high-performance and low-cost alloys, which were experimentally validated with 90% fitting accuracy<sup>[33]</sup>. Jiang *et al.* developed a MLDS consisting of both component design (P2C) and performance prediction (C2P) components, and identified the artificial neural network (ANN) model as the most suitable<sup>[34]</sup>. The MLDS is used to design alloy components that meet the target properties. The experimental results show that the designed alloy has excellent properties, which proves the reliability of the system. It embodies the advantages of data-driven and performance-oriented ML component design method compared with traditional trial-and-error method in design efficiency and performance optimisation, and provides theoretical basis and practical guidance for intelligent design of aluminium alloys. In summary, big data-driven ML methods are an effective approach to exploring the “strength-ductility trade-off” dilemma of Al-Zn-Mg-Cu alloys.

In the field of aluminium alloy material research, past research achievements have been abundant, laying a solid foundation for our work. However, an undeniable fact is that most of these studies tend to use relatively common algorithm models in method selection. Although these traditional algorithm models have provided some valuable ideas and results for the design of aluminium alloys to a certain extent, they still face enormous challenges in solving the key problem in alloy performance optimisation - the “strength-ductility trade-off”<sup>[35]</sup>. This dilemma has limited the performance improvement of aluminium alloys in practical applications to a certain extent and prompted researchers to continuously explore more innovative and effective methods to overcome this problem in order to meet the growing engineering application needs and higher material performance standards. Moreover, previous studies may have focused on the optimisation of a certain specific performance. Therefore, starting from the actual application scenario and considering the usage requirements of aluminium alloys in different fields, we focus on the balanced optimisation of strength and EL.

In view of the challenges still faced by existing studies in optimising the “strength-ductility trade-off” dilemma for high-strength Al-Zn-Mg-Cu alloys, this study aims to explore alloy composition design by introducing ML models. A database containing alloy components and corresponding mechanical properties is constructed, and a multi-algorithm integrated ML approach is used to accelerate the screening of alloys with the goal of achieving the desired mechanical properties. The integrated strategy integrates the advantages of multiple models, which can significantly improve the accuracy and reliability of the prediction of alloy properties and provide a more effective tool for alloy design. During the experimental validation process, the effects of extrusion (E) and rolling (R) processes on the properties of the designed alloys are further compared, and the key role of process factors in regulating alloy properties is discussed in depth. This study provides new ideas for solving the “strength-ductility trade-off”, which can help develop more competitive aluminium alloys and promote their widespread use in engineering applications.

## MATERIALS AND METHODS

### ML methods

Scikit-learn in Python is a common choice for algorithmic implementations due to its extensive set of ML algorithms, ease of use, and large community support. Compared with Matlab or R language, it has a simple and intuitive application program interface and does not rely on large computing power, so that both beginners and experienced developers can easily get started, with strong realisable and low cost. Therefore, this study is based on python implementation. The dataset plays a key role in the performance of ML models. The Al-Zn-Mg-Cu alloy data used in this study was obtained from relevant published literature. To ensure that the data is as accurate as possible and that the results of composition and properties are not influenced by the preparation process<sup>[36,37]</sup>, the data collected from the published literature was explicitly restricted to the alloy preparation processes, including melt casting, homogenisation, conventional plastic deformation, conventional solid solution treatment, and single-stage ageing treatment (T6 heat treatment process). The screening was initially based on about 160 publications. After careful evaluation and screening of the remaining literature of about 48 publications, 147 sets of high-quality data met the above criteria and were collected to create a dataset of the composition and mechanical properties of Al-Zn-Mg-Cu alloys. The aluminium alloy dataset established can be found in the [Supplementary Materials](#). Although these 147 data sets may be relatively limited in number, each contains a rich set of feature dimensions that are sufficient to ensure that the model can capture trends in the data. Advanced data analysis methods and model evaluation metrics were used during the study to ensure the validity and representativeness of these data, and the training results were validated and optimised several times to ensure the stability and generalisation ability of the model.

The initial data set is cleaned and divided and the outliers and missing values in the dataset were statistics. The Sample with more individual abnormal outliers and missing values were excluded, and the missing values were filled in using mode to minimise their effects and avoid serious bias on the prediction results. The feature values of input alloying elements were standardised to scale the features into a standard normal distribution. The output target values, ultimate tensile strength (UTS), were added with Gaussian noise to prevent overfitting. In Gaussian noise, two key parameters, the mean ( $\mu$ ) and standard deviation ( $\sigma$ ), are used to describe the statistical characteristics of the noise. The alloying elements in the final processed dataset primarily include 14 kinds (Zn, Mg, Cu, Zr, Mn, Ti, Sn, Cr, Sc, Ce, Ni, Y, Er, and Sr). The mechanical property parameters mainly contain UTS, YS, and EL. [Table 1](#) shows the range of alloy components and properties (RE stands for rare earth elements) in the Al-Zn-Mg-Cu alloy dataset of this study. After the data cleaning was completed, Pearson's feature correlation analysis was performed to reveal the association of the features with the target variables.

During the model construction stage, various ML algorithms were selected for UTS single-objective optimisation, including linear regression (LR), L1 regularisation (L1), ANN, K-nearest neighbours (KNN), decision trees (DT), support vector machines (SVM), random forest (RF), and multi-algorithmic integration models of bagging, boosting, stacking, and voting. Then, selecting the model with the best evaluation results as the final ML model and the forward multi-objective screening combined with the reverse design strategy was used to design the alloy components with desirable properties.

For better evaluation of the model's generalisation capability, a random division is used to divide the dataset randomly by 10%-20% as the test set, and the remaining 80%-90% as the training set, which is repeated 100 times. The average of prediction results is taken as the result, to reduce the fluctuation of prediction caused by the data differences in the division of the data. The  $R^2$  is used to evaluate the model.

**Table 1. Range of alloy composition and performance in the Al-Zn-Mg-Cu alloy dataset**

Element composition (wt.%)														UTS (MPa)	YS (MPa)	EL (%)
Zn	Mg	Cu	Zr	Mn	Ti	Sn	Cr	Sc	Ce	Ni	Y	Er	Sr			
0-11.30	0-5.64	0-4.14	0-0.51	0-0.46	0-0.12	0-0.51	0-0.28	0-0.50	0-0.30	0-0.57	0-0.32	0-0.13	0-0.02	276-812.40	214-752.90	0.3-21.35

UTS: Ultimate tensile strength; YS: yield strength; EL: elongation.

## Experimental

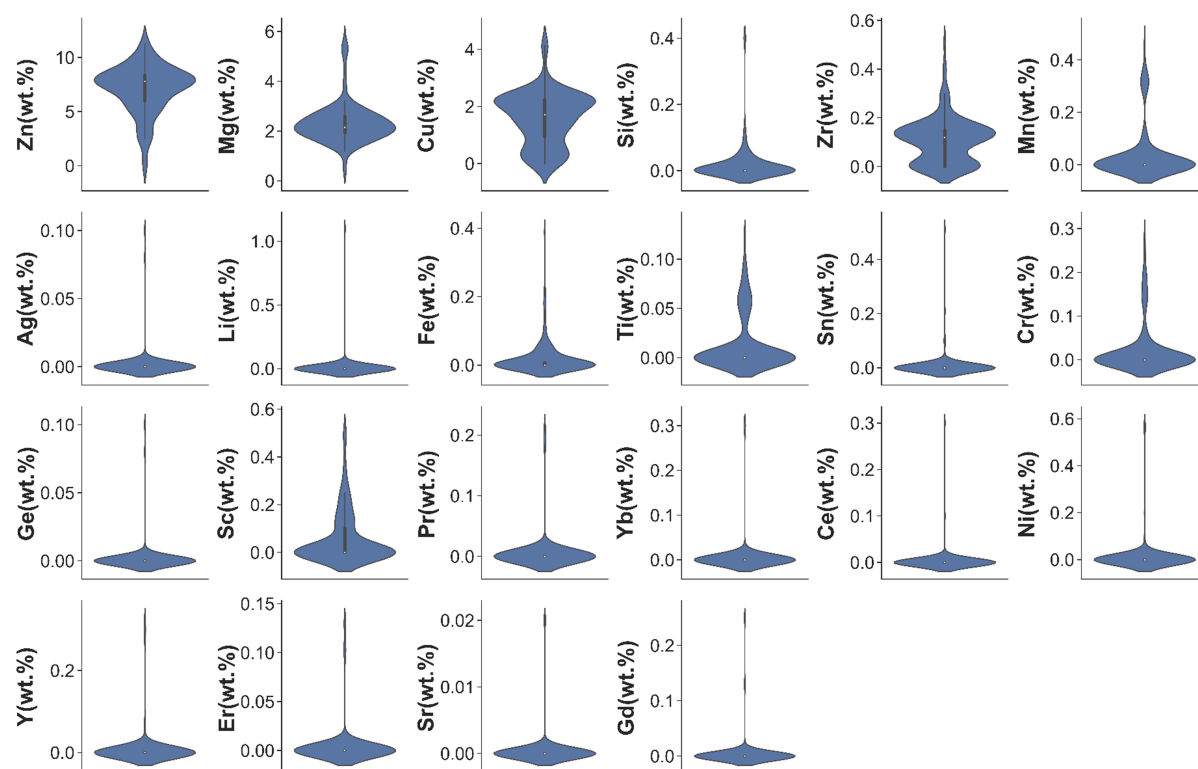
Industrial pure Al (99.70%, purity), pure Mg (99.97%), pure Zn (99.97%), pure Sn (99.99%), Al-50%Cu, Al-5%Zr, Al-10%Mn, Al-10%Ti, Al-2%Sc, Al-10%Ce, Al-10%Ni, Al-5%Y, and Al-5%Cr master alloys (compositions in wt.%) were used as experimental raw materials for the experiment. The mould is preheated before casting to avoid a sharp drop in the temperature of the molten metal during the pouring process. In the alloy casting process, a bottom pouring method is used for pouring, and a metal mould is used. The specific dimensions are shown in the [Supplementary Figure 1](#). During the pouring process, a uniform pouring speed is maintained to make the molten metal fill the mold smoothly and reduce the formation of bubbles and inclusions, thereby improving the performance and surface quality of the casting.  $C_2Cl_6$  was used to refine and degas during the melting process, and pouring was carried out at 710 °C to obtain alloy ingots. Homogenisation was carried out for 440 °C @ 24 h + 470 °C @ 24 h. Extrusion was conducted by a four-column hydraulic press (YQ32-100) at 400 °C with an extrusion ratio of about 6. Rolling was performed by a two-roll multifunctional experimental mill with a total reduction of about 70%. Subsequently, the solution treatment was executed at 450 °C @ 1 h + 470 °C @ 1 h + 480 °C @ 1 h, followed by aging at 120 °C for 24 h. Finally, three groups of samples were obtained through the final complete heat treatment, including A1-A4 without deformation, E1-E4 after extrusion deformation, and R1-R4 by rolling process.

The metallographic detection surface was etched using a Keller etchant (5 mL  $HNO_3$ , 3 mL HCl, 2 mL HF, and 190 mL distilled water) for 7-9 s and the microstructures of the different experimental samples were observed using optical microscopy (Nikon, MA200). To observe the morphology and composition of phases, scanning electron microscopy (SEM; Zeiss, EVO 18) and energy dispersive X-ray spectroscopy (EDS; Bruker, X Flash Detector 6|30) were used. The phase composition and crystal structure were measured using X-ray diffraction (XRD; Rigaku, Mini Flex 600) in the scan range of 10°-90°, and scan speed of 5°/min. Tensile tests were performed at room temperature on an MTS-E45.305 universal testing machine using sheet specimens (as described in Ref. <sup>[38]</sup>), and the tensile rate was set to 1 mm/min based on the national standard metal material tensile test method. Notably, the average values of the parameters of the three alloys under the same conditions are considered for the experimental test data to ensure the accuracy of the results.

## RESULTS AND DISCUSSION

### Construction of ML model

Related research<sup>[39]</sup> has demonstrated that the UTS is less sensitive to specimen size, surface accuracy, preparation parameters and data post-processing; thus, it can provide more accurate and stable measurements. Based on this factor, the UTS is used as the primary output performance target and the collected compositional data as descriptors to conduct single-objective optimisation and explore the optimal ML screening model for Al-Zn-Mg-Cu alloys. The 80% of the data was chosen as the training set and the remaining 20% as the test set. The overall distribution of the initial untreated dataset is illustrated in [Figure 1](#).

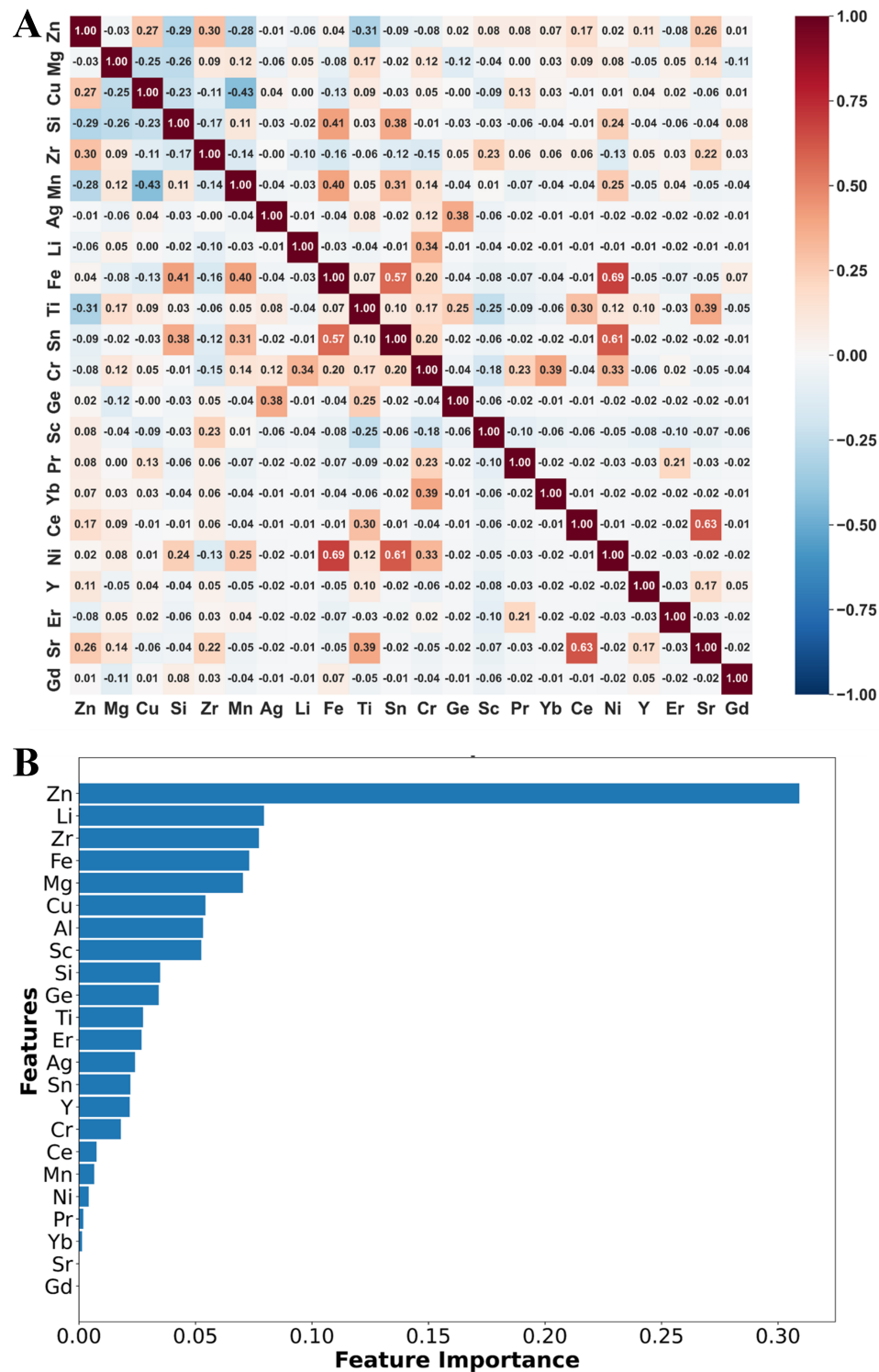


**Figure 1.** Distribution of alloy element feature values in the initial untreated dataset.

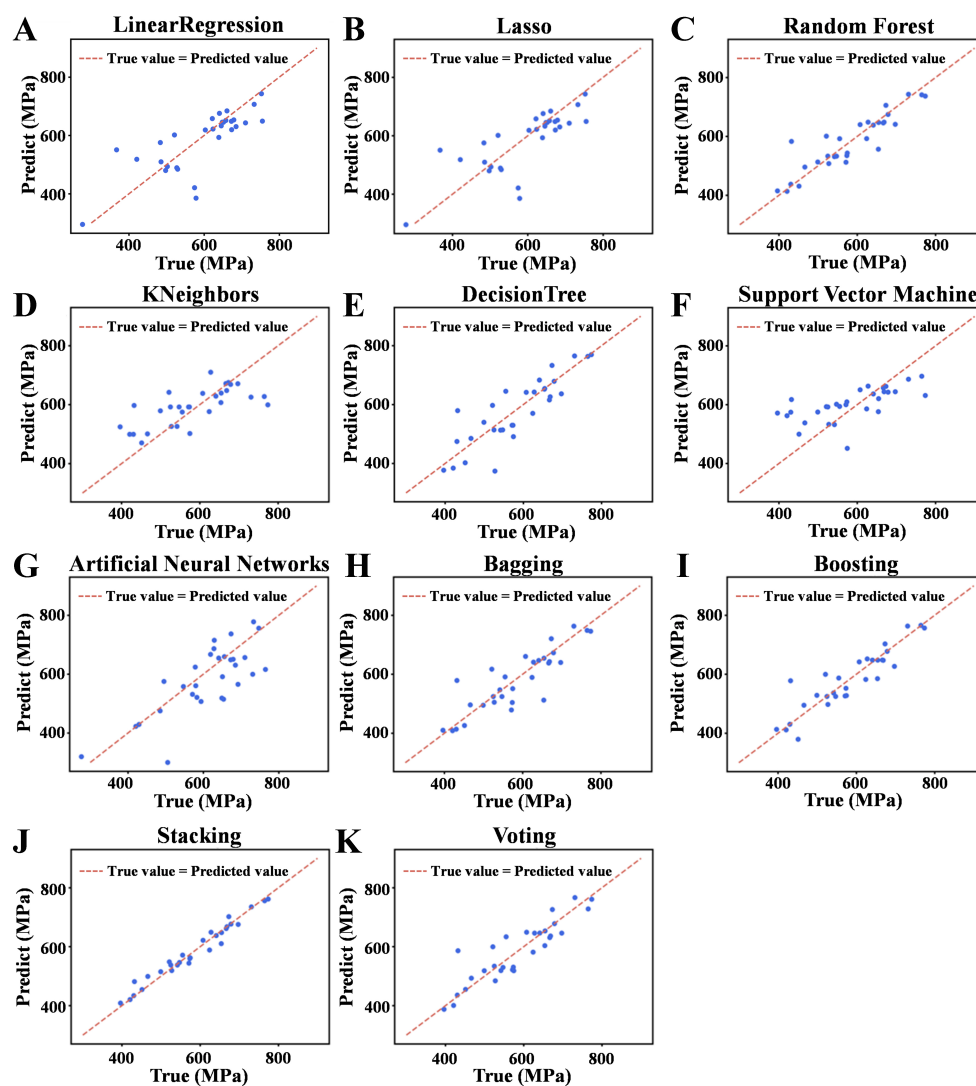
To design an accurate ML predictive model, it is essential to select appropriate features that are directly related to the target properties. In the modelling process, feature analysis helps to understand the extent to which a feature affects the target variable. This can be used to select the most relevant features, eliminate irrelevant or redundant features, reduce the risk of overfitting, and enhance the model's generalisation capability. After that, features with excessive correlation are removed, followed by feature correlation analysis and ranking of importance, as shown in [Figure 2](#).

The feature Pearson correlation heatmap is shown in [Figure 2A](#), with a value range between  $[-1, 1]$ , where  $-1$  represents a complete negative correlation,  $+1$  indicates a complete positive correlation, and  $0$  means no correlation. The depth of the colour in the heatmap directly reflects the degree of correlation between the features, with darker colours indicating a more significant correlation. It can be observed that the correlation of the features is all less than  $0.8$  after treatment, indicating no strong correlation between them, and the features will not affect each other, thus reducing the error of the results.

The demonstration of feature importance is crucial to our understanding of the decision-making process of the model and the relative contribution of the variables. More emphasis is placed on revealing the impact of variables from the perspective of model prediction. It considers not only the direct relationship between variables, but also their combined contribution to the prediction goal during the model learning process, and not only in direct relation to the content of the components. For elements with varying content, low feature importance does not mean that they are irrelevant for performance. This model-learning-based assessment gives us a deeper understanding of which factors are actually driving alloy performance, rather than just superficial data distribution features. Here, an impurity-based built-in feature importance method was employed to assess feature significance. Specifically, we used the extreme gradient boosting random



forest regressor (XGBRFRegressor) from the extreme gradient boosting (XGBoost) library for regression tasks. By training the model on the feature data, we obtained the importance scores for each feature using the model's `feature_importances` attribute. This method calculates the contribution of each feature during

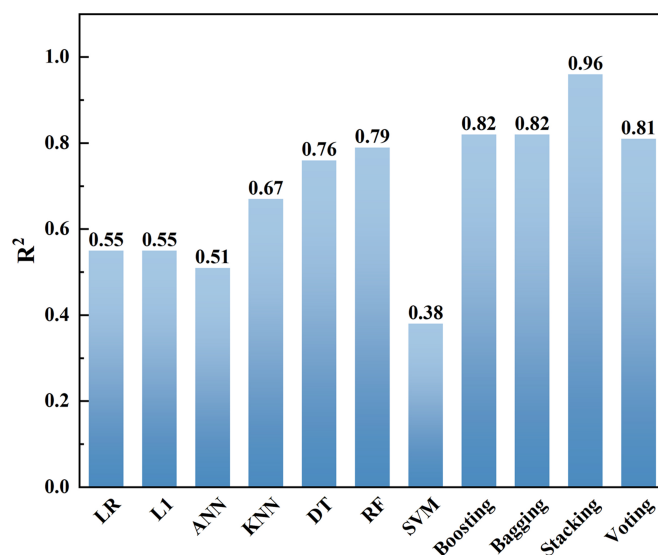


**Figure 3.** Scatter plot of UTS predictions by different ML models on the test set. (A) LR; (B) Lasso; (C) RF; (D) KNeighbors; (E) DT; (F) SVM; (G) ANNs; (H) Bagging; (I) Boosting; (J) Stacking; (K) Voting. UTS: Ultimate tensile strength; ML: machine learning; LR: linear regression; RF: random forest; DT: decision trees; SVM: support vector machines; ANNs: artificial neural networks.

tree splits, thereby reflecting its influence on the model's predictions. Organising the features and their importance scores into a data frame, followed by sorting and visualisation, allows us to clearly identify the features that have the greatest impact on the target variable. This approach effectively handles high-dimensional datasets and provides a solid foundation for feature selection. As can be seen from the feature importance ranking in Figure 2B, most of the features influence the results within the normal range of 0.3; however, the Zn element has the greatest effect on the target performance of UTS compared with the other alloying elements. Individual alloying elements Sr and Gd had almost no effect on the target values, so the data of these alloying elements were removed from the dataset.

The scatter plots of the predicted results of various models in the test are presented in Figure 3. Analysis results show the  $R^2$  of various models, as presented in Figure 4. However, the SVM model is unsuitable for this dataset with an  $R^2$  of only 0.38. In contrast, several ensemble learning models show better accuracy with  $R^2$  greater than 0.8. This is because an ensemble model can overcome the limitations of individual models





**Figure 4.** Prediction results  $R^2$  of various ML models in test.

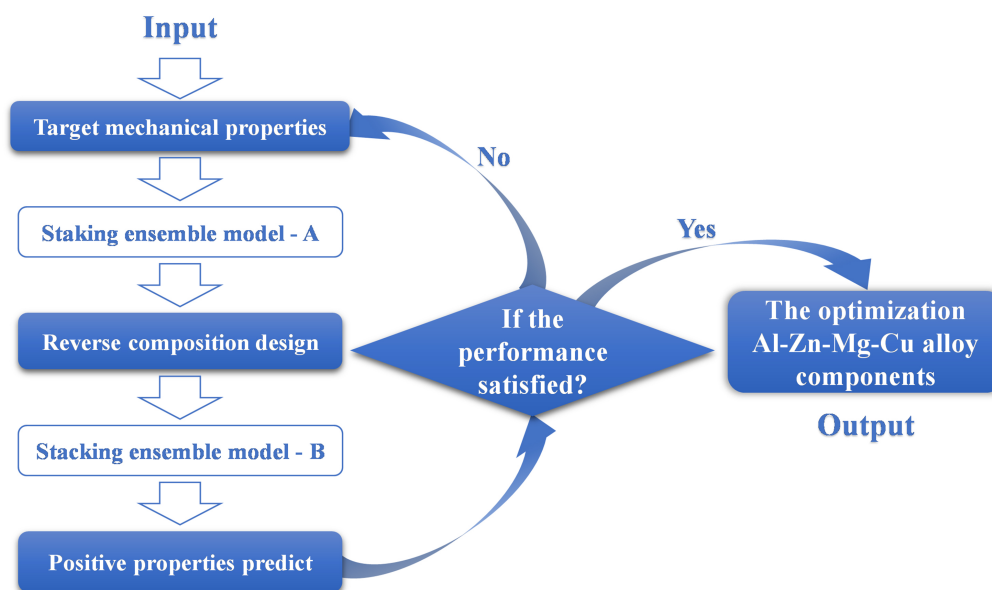
and exhibit more robust predictions in different situations. They can integrate the advantages of multiple models, reduce the risk of overfitting and improve the generalisation capabilities of models.

The stacking model is the most effective among the ensemble models with an  $R^2$  of 0.96 which can also be seen from the scatter plot. Consequently, the stacking ensemble model was selected as the final alloy composition screening model. In this model, a heterogeneous, sequential integration of KNN, RF and ANN algorithms is used as weak learners in order, with the RF algorithm serving as the meta-model to explore the best model parameters and finally get the model. Here, the ML model that correlates composition with performance was established and the prediction accuracy of the model is significantly improved by using stacking ensemble learning model.

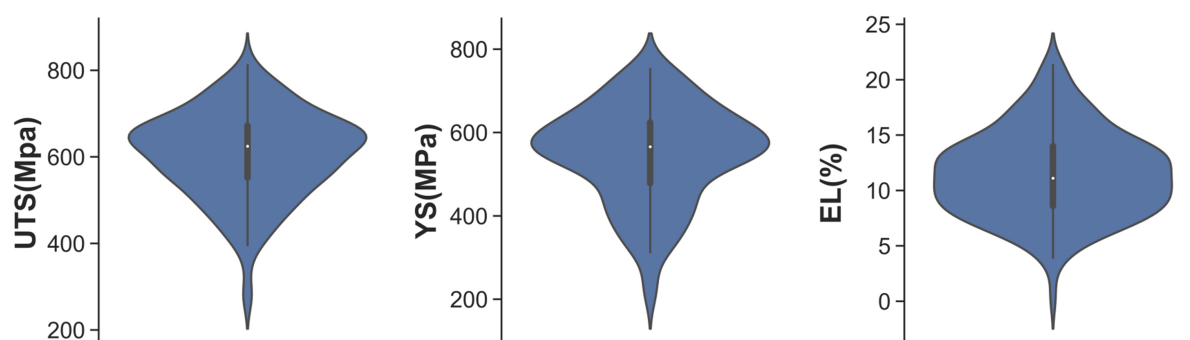
The final output is influenced by a variety of factors. Firstly, the accuracy and diversity of the basis models are crucial. Additionally, it is important for the meta-learner to effectively integrate the outputs of the base models to capture the potential relationships for improving the prediction accuracy and stability. Furthermore, data quality and feature engineering indirectly affect the output, and good quality data and reasonable feature selection can improve the learning effect of the model and enhance the reliability of the results.

### Construction of MLDS

The forward propagation task is crucial in alloy design strategies. Models are trained based on a large amount of known data and learn complex mapping relationships between alloy components and mechanical properties. When new property data are provided, the model can infer the most likely alloy component combinations based on the already learnt mapping relationships through forward propagation, and further predict the possible mechanical properties. The input mechanical property data is processed layer by layer through this process, iterating through forward and backward propagation. The predicted properties are compared with the desired target properties to ensure they satisfy the set error requirements. If the requirements are not met, the design prediction cycle is repeated until satisfactory composition and mechanical properties are obtained. The model is continuously optimised to improve the prediction accuracy. The specific flow chart is shown in [Figure 5](#).



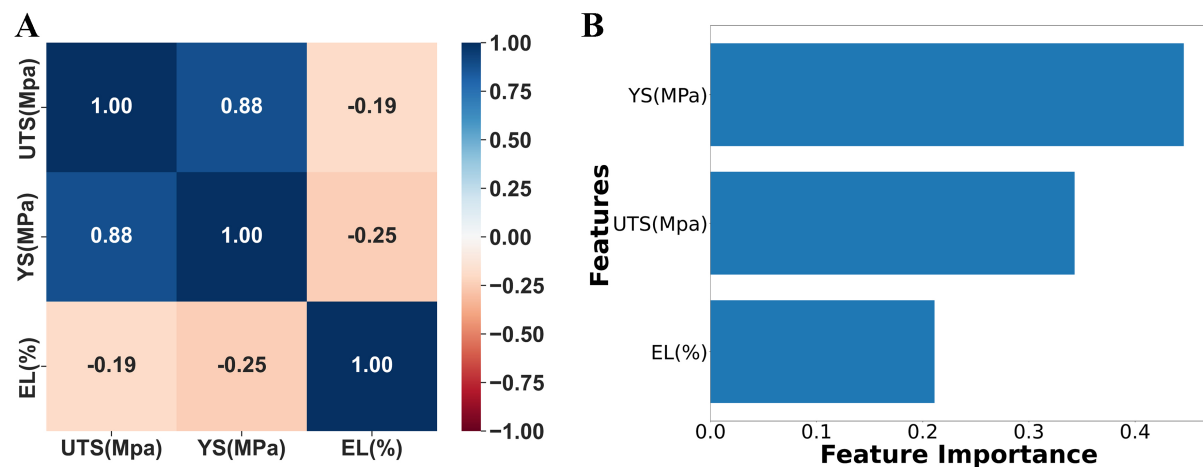
**Figure 5.** Schematic diagram of target properties screening process of Al-Zn-Mg-Cu alloy.



**Figure 6.** Distribution of feature values of mechanical properties parameters.

The mechanical property parameters are used as descriptors in the reverse design of the alloy composition to output the alloy component content. Similar to the previous ML steps, the data was pre-processed initially to view the distribution of the feature values of the mechanical property parameters, as shown in Figure 6. Based on this, the features were analysed, as shown in Figure 7. From the Pearson correlation heatmap [Figure 7A], the UTS and YS are shown as a strong positive correlation from the analysis of a large amount of data, while EL shows a negative correlation with strength. From the results of the feature importance ranking [Figure 7B], the YS has the greatest influence on the alloy component reverse design results.

After completing the necessary preliminary work, the alloy component design is carried out. From the desired mechanical properties, the optimal material composition is deduced backwards. The main alloying elements Zn, Mg and Cu, along with the more widely studied elements Zr, Ti, Sn and Sc, have a richer accumulation of relevant experimental data, and thus their data points are more densely populated in our dataset. In contrast, other alloying elements are less frequently used or less studied, resulting in a relatively limited amount of data. This is a common phenomenon in the field of aluminium alloy research and an

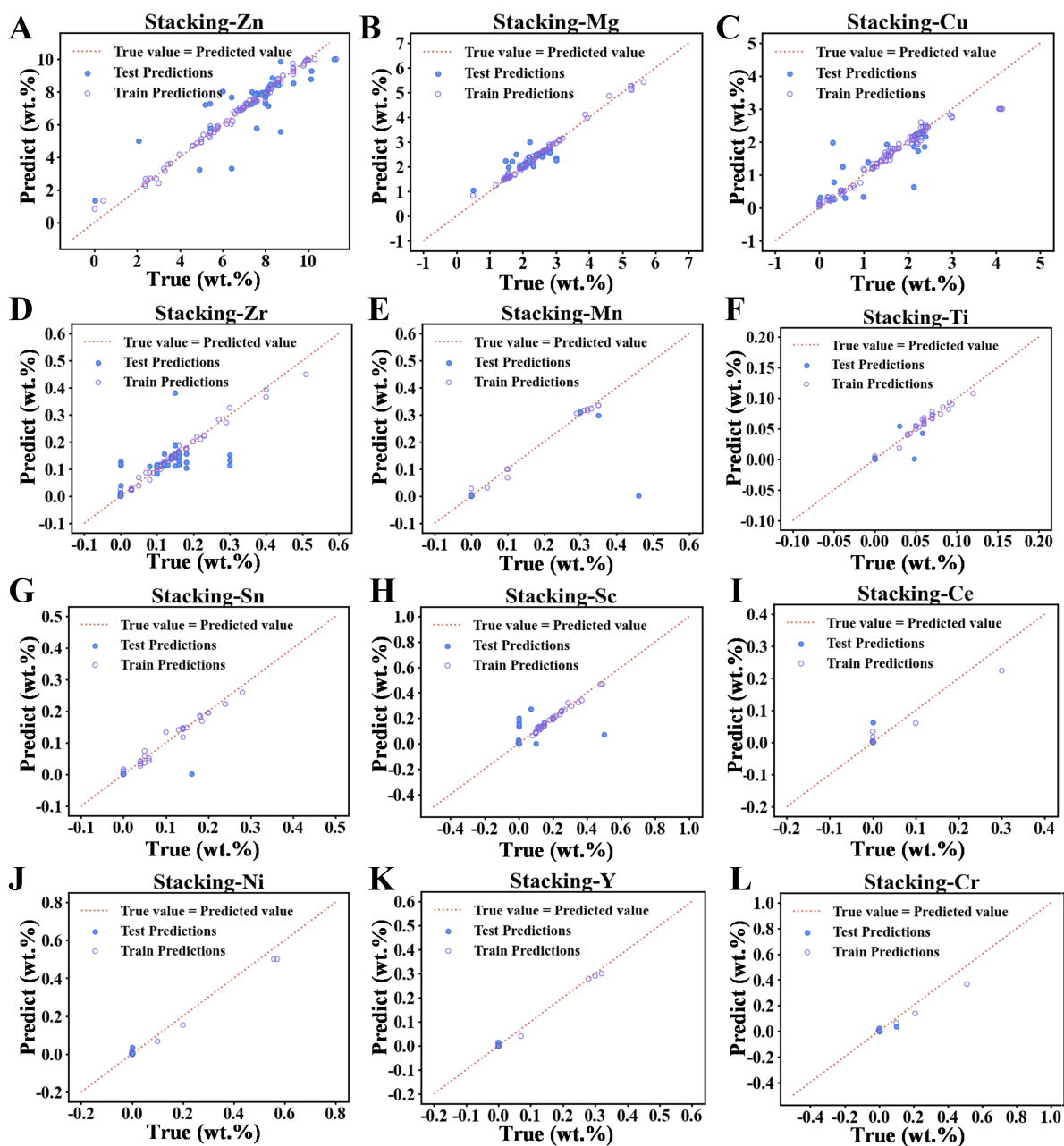


**Figure 7.** Pearson feature correlation analysis between mechanical properties parameters. (A) heatmap; (B) feature importance ranking.

objective challenge in data collection. However, ML models are equipped with the ability to deal with data imbalance and make full use of the data of different elements through integrated learning methods, which can be useful even for elements with less data. In addition, feature combinations that contribute the most to performance prediction are filtered by evaluating feature importance during model training, which improves the learning efficiency and prediction accuracy of the model with limited data. To validate the effectiveness of the model, we conducted a rigorous model evaluation.

The results are presented as a scattered plot [Figure 8] where the data from the training set and the test set are clearly distinguished. Due to the small amount of data available for some alloy compositions, 90% of the data was chosen as the training set and the remaining 10% as the test set, which was to reduce possible prediction errors due to insufficient data. During the backward design process, Zn element has an  $R^2$  of 0.98 in the training set, indicating that the model fits the training data very well. However, it scores 0.62 in the test set, suggesting that the model's generalisation ability on the unknown data has been reduced. The situation is similar for elemental magnesium and copper, whose ratings in the training set are 0.99 and 0.95, respectively, while in the test set, they are 0.45 and 0.38, respectively. In the forecasting process, the prediction results of each element are not independent of each other. We consider the interactions and correlations between the elements. In the process of model training, the model can learn these complex relationships by introducing multiple features and considering the correlation constraints between elements.

Based on the alloy element compositions generated by the previous design, the obtained alloy compositions will be forward-predicted to the corresponding alloy mechanical properties. The forecasted properties will be compared with the target properties to check whether the predefined error tolerances are met. If not, the cycle of design prediction will be repeated until a satisfactory alloy composition and mechanical properties are achieved. Upon re-examining the features within the dataset, the Pearson correlation heatmap and the feature importance ranking are shown in Figure 9. It is observed that the intercorrelation among features is negligible, thereby precluding any impact on the results from feature interdependencies. In addition, there are no features in the feature importance ranking that have too much or no effect on the results. After the feature analysis, the constructed model was employed for multi-objective forward performance predictions. The scatter plot in Figure 10 illustrates that the predictive outcomes for various mechanical performance parameters are relatively accurate.



**Figure 8.** Scatter plot of the prediction results of the reverse component design in test. (A)Zn; (B)Mg; (C)Cu; (D)Zr; (E)Mn; (F)Ti; (G)Sn; (H)Sc; (I)Ce; (J)Ni; (K)Y; (L)Cr.

Utilising the stacking ensemble model, we engaged in the design of alloy compositions. The correlation analysis depicted in Figure 7 indicates a robust positive correlation between UTS and YS, while an EL is inversely related to strength. Consequently, the input target performance parameters are detailed in Table 2. We have designed four sets of target performance alloys labelled A1-A4, aspiring to develop Al-Zn-Mg-Cu alloys with both superior strength and ductility.

By inputting the desired sets of target mechanical properties of both high strength and ductility Al-Zn-Mg-Cu alloys, four groups of alloy components were obtained based on the stacking model, and the design

Table 2. Target performance values adopted for alloy component design

	UTS (MPa)	YS (MPa)	EL (%)
A1	710.0	670.0	9.5
A2	720.0	690.0	9.0
A3	730.0	710.0	8.5
A4	740.0	730.0	8.0

UTS: Ultimate tensile strength; YS: yield strength; EL: elongation.

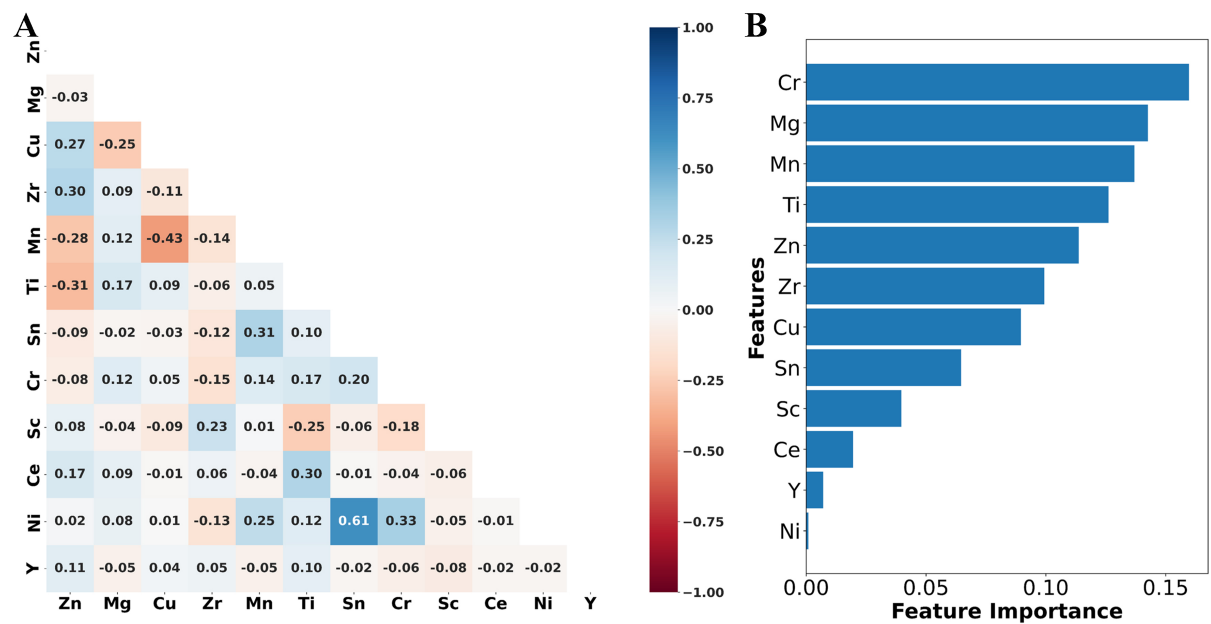


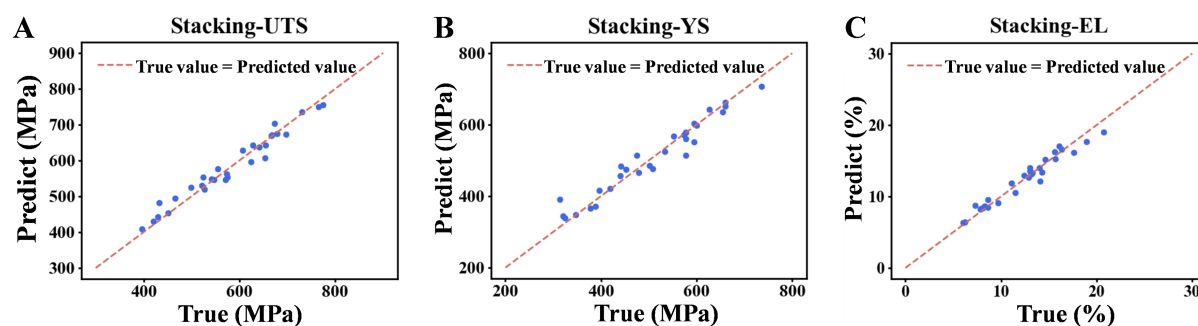
Figure 9. Pearson feature correlation analysis of multi-target positive prediction. (A) feature correlation heatmap; (B) feature importance ranking.

contents of different alloying elements are shown in Table 3.

The elements Ti, Y and Ce significantly increase the strength of the alloys through grain refinement and the formation of reinforcing phases ( $\text{TiAl}_3$ ,  $\text{Al}_2\text{Y}$ ,  $\text{CeAl}_4$ , etc.). The grain refinement effect of Ti and the improvement of the grain boundary structure by moderate amounts of Y and Ce are beneficial to the ductility. The strength of aluminium alloys containing Sc tends to increase more than that of normal aluminium alloys due to the function of the  $\text{Al}_3\text{Sc}$  phase, and the formation of intermetallic compounds by Sn and Ni also contributes to a certain extent to the enhancement of the strength, but the content needs to be controlled to avoid a negative effect on the toughness. The formation of undesirable phases due to high levels of Sn and Ni and the addition of excessive amounts of elements such as Sc may reduce toughness. Sc is a relatively expensive element and its high cost limits its widespread use in large-scale industrial production. In alloy design and production, a balance must be struck between the improved performance of these elements and their increased cost. Over-reliance on expensive elements may make the alloy too costly and affect its competitiveness in the marketplace, while reducing the amount of these elements added may not satisfy the performance requirements.

**Table 3.** Four groups of Al-Zn-Mg-Cu alloy components that are designed with the stacking model constructed based on Figure 5 (wt.%)

	Zn	Mg	Cu	Zr	Mn	Ti	Sn	Ni	Cr	Sc	Y	Ce	Al
<b>A1</b>	8.520	2.959	1.935	0.146	0.001	0.001	0.015	0.001	0.001	0.001	0.001	0.001	Bal.
<b>A2</b>	8.946	2.543	1.229	0.075	0.001	0.001	0.100	0.001	0.001	0.001	0.001	0.001	Bal.
<b>A3</b>	9.226	2.129	0.637	0.076	0.001	0.046	0.001	0.001	0.001	0.001	0.010	0.031	Bal.
<b>A4</b>	9.280	2.440	0.607	0.004	0.001	0.001	0.138	0.055	0.001	0.218	0.001	0.001	Bal.

**Figure 10.** Scatter plot of multi-objective forward performance prediction results in test. (A) UTS; (B) YS; (C) EL. UTS: Ultimate tensile strength; YS: yield strength; EL: elongation.

### Composition and microstructure of the experimentally prepared alloys

Four groups of alloys designed by ML were melted and cast. To verify that the actual chemical compositions of the alloys are consistent with the design, the actual chemical compositions of the prepared alloys were examined using inductively coupled plasma atomic emission spectroscopy (ICP-AES) as shown in Table 4. The results indicate that the prepared alloy components are basically consistent with the compositional contents designed by ML.

As shown in Figure 11, XRD analysis of the phase composition within the four sets of alloys was conducted, with the XRD spectra exhibiting the characteristic diffraction peaks of  $\alpha$ -Al,  $\text{Al}_2\text{CuMg}$ , T-AlZnMgCu,  $\text{Al}_2\text{Cu}$ , and  $\text{MgZn}_2$  phases.  $\alpha$ -Al constitutes the principal matrix phase of the aluminium alloy, and  $\text{Al}_2\text{CuMg}$ , T-AlZnMgCu,  $\text{Al}_2\text{Cu}$ , and  $\text{MgZn}_2$  are common strengthening phases. The existence of these phases exerts a significant impact on the hardness, strength, and ductility of the alloy.

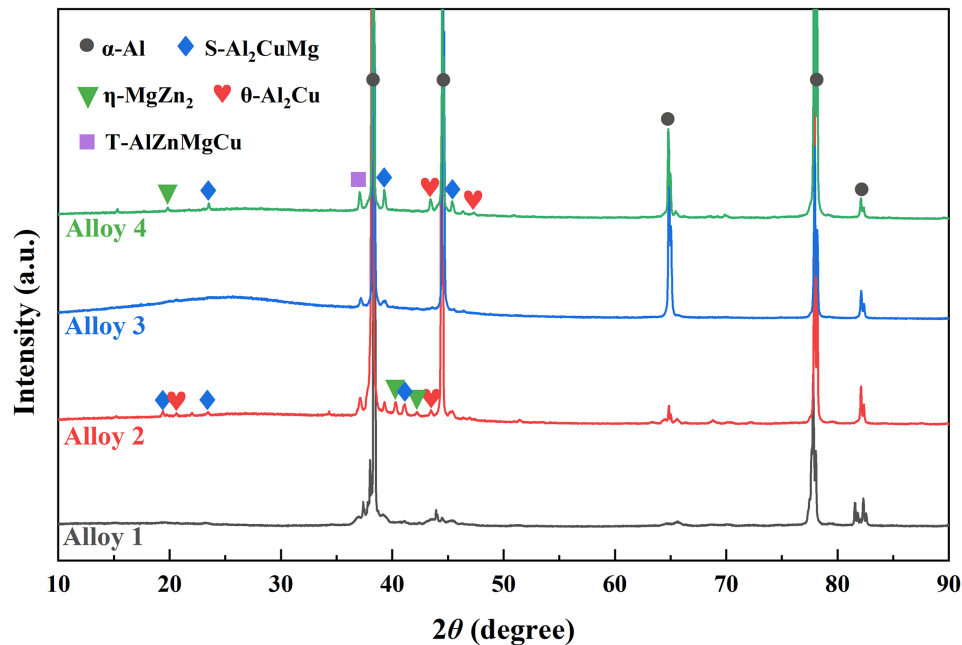
Figure 12 shows the SEM microstructure and EDS analysis of the second phase. In the microstructure of alloy A1, point 1 is the T-AlZnMgCu phase, while in alloy A2, point 2 is the  $\text{Al}_2\text{CuMg}$  phase. In alloy A3, the irregularly shaped, bright white phase is the  $\text{Al}_2\text{Cu}$  phase, and it is evident that many black, punctate  $\text{MgZn}_2$  phases have formed in alloy A4. However, this phenomenon was not observed in the other alloys. Supplementary Table 1 displays the EDS point scanning test data for the different phases of the alloy, as shown in Figure 12. The data confirm the chemical compositions of the respective phases, which consist of  $\text{Al}_2\text{CuMg}$ , T-AlZnMgCu,  $\text{Al}_2\text{Cu}$ , and  $\text{MgZn}_2$ .

So far, we mainly constructed a model between composition and alloy properties by ML methods, and quantitatively analysed these properties under fixed process conditions. However, the influence of deformation processes on alloy properties, such as extrusion and rolling deformation processes, is also crucial. Therefore, this paper introduces an experimental comparative study between extrusion and rolling processes to analyse the effects of different deformation processes on alloy properties and provide a



**Table 4. Actual elements content of four groups of Al-Zn-Mg-Cu alloys (wt.%)**

	Zn	Mg	Cu	Zr	Mn	Ti	Sn	Ni	Cr	Sc	Y	Ce	Al
<b>A1</b>	8.670	3.230	2.280	0.110	0.003	0.004	0.011	0.006	0.002	0.001	0.001	0.002	Bal.
<b>A2</b>	9.110	2.960	1.400	0.060	0.003	0.003	0.056	0.005	0.003	0.001	0.001	0.001	Bal.
<b>A3</b>	9.440	2.700	0.730	0.060	0.003	0.027	0.002	0.006	0.002	0.001	0.003	0.013	Bal.
<b>A4</b>	9.970	2.740	0.690	0.004	0.002	0.004	0.094	0.053	0.003	0.180	0.001	0.001	Bal.

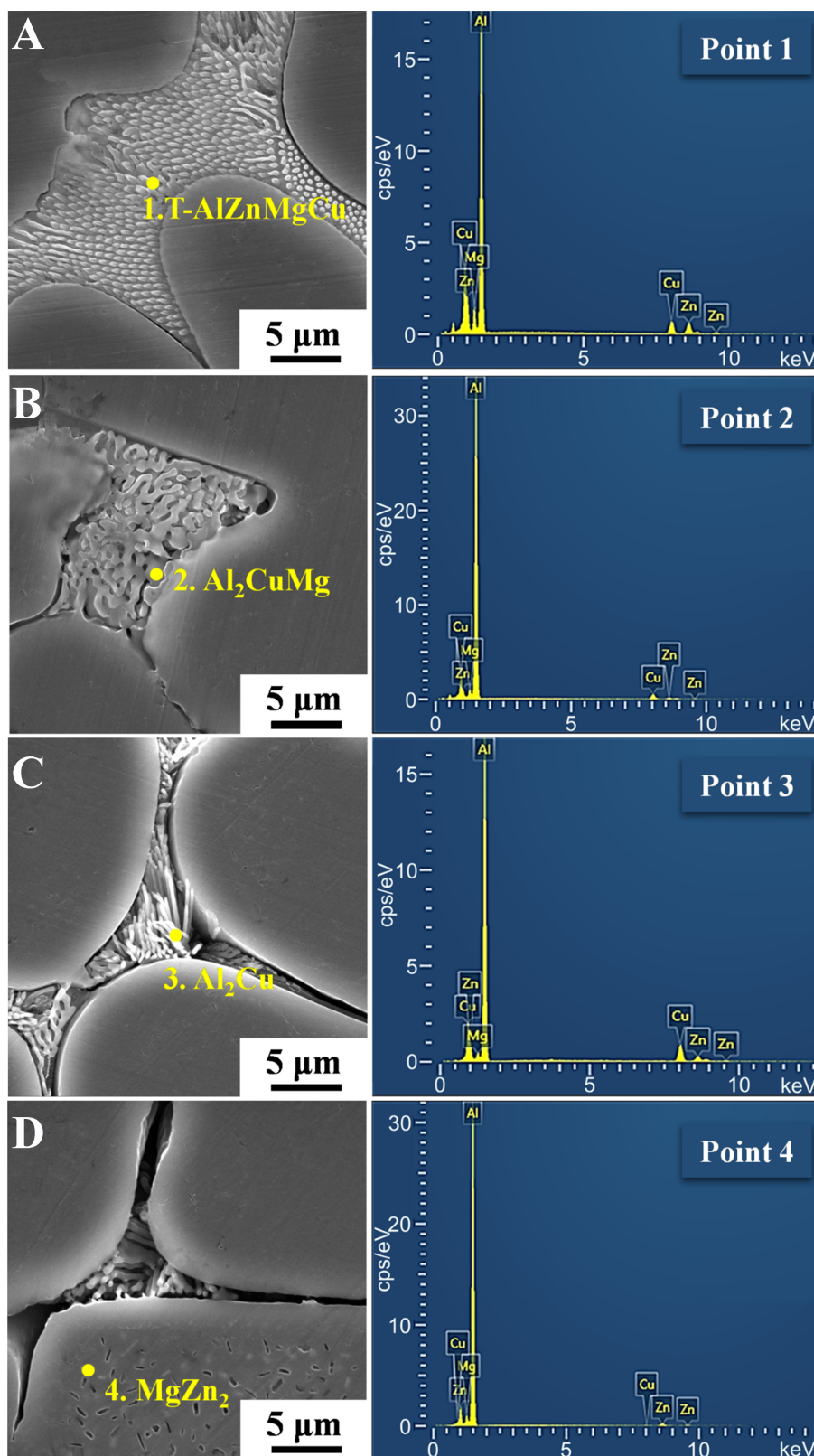
**Figure 11.** XRD phase analysis of four groups of Al-Zn-Mg-Cu alloys. XRD: X-ray diffraction.

reference basis for process optimisation in industrial production. Here, the alloy is further homogenised, deformed, solution-aged and heat-treated. The two deformation processes of extrusion (E) and rolling (R) are used here, while the heat treatment process and parameters remain unchanged.

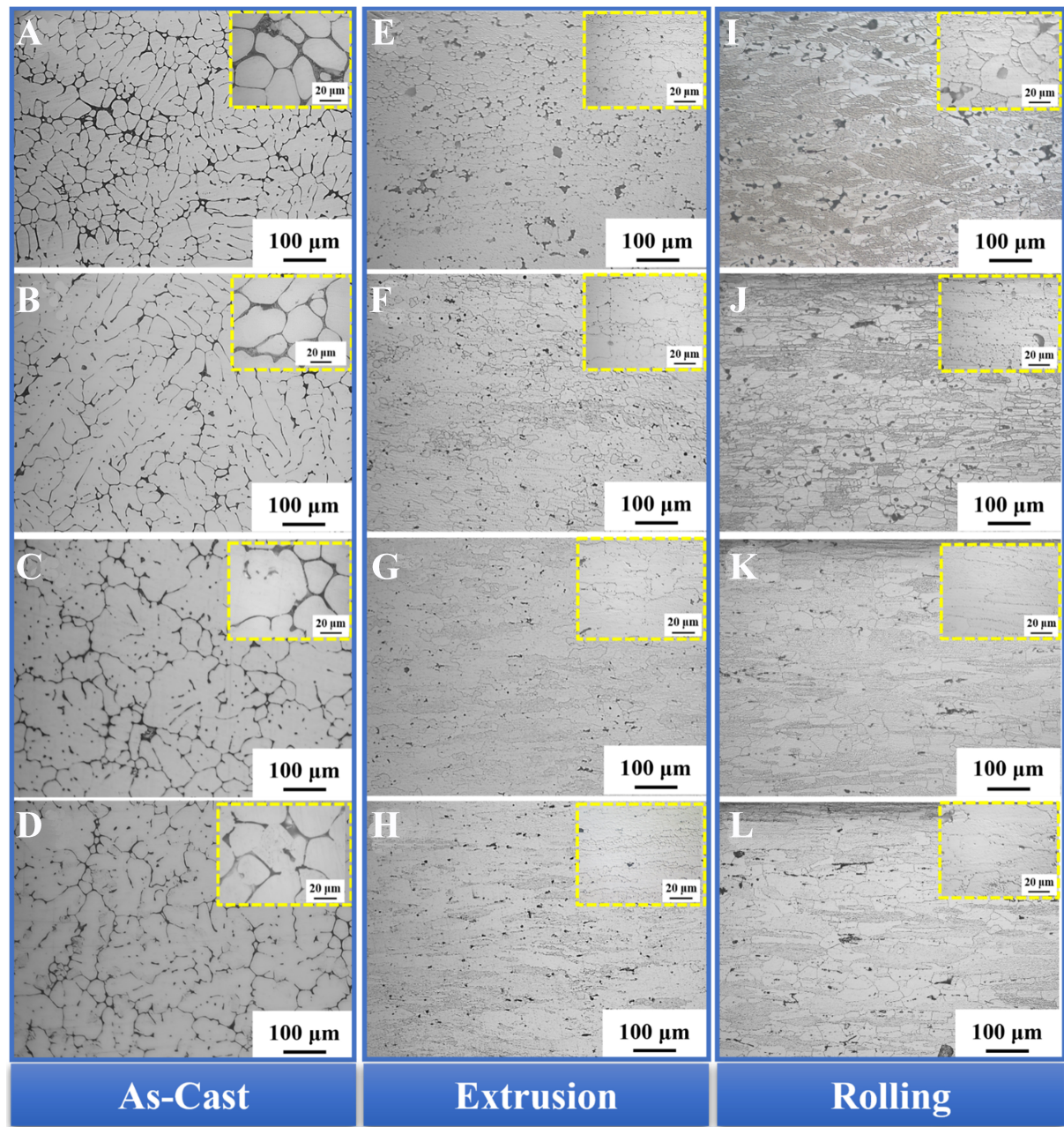
To observe the microstructures of the alloys, Figure 13 illustrates the microstructures of four groups of Al-Zn-Mg-Cu alloys under different treatment conditions. The microstructures of the as-cast samples without solid solution and ageing treatment are shown in Figure 13A-D, while those of the extruded and rolled alloys after solid solution treatment are presented in Figure 13E-H and 13I-L, respectively. The microstructure of the alloy predominantly contains the white dendritic  $\alpha$ -Al matrix and the black coral-like eutectic structures. The grey coral-like regions are eutectic phases formed during the non-equilibrium solidification; there is no pronounced directionality. A comparison of the metallographic structures with varying elemental compositions reveals differences in the morphology and size of the  $\alpha$ -Al grains.

### Mechanical properties of different processes

The stress-strain curves of the mechanical properties and the strength EL values obtained are shown in Figure 14. It can be observed that four different alloy samples E1-E4 obtained after heat treatment of the extrusion process show excellent mechanical properties, which is consistent with the trend of alloy properties designed by ML. This shows that the alloy properties predicted by the ML method are in good



**Figure 12.** SEM microstructure and EDS analysis of the second phase. (A-D): A1-A4. SEM: Scanning electron microscopy; EDS: energy dispersive X-ray spectroscopy.

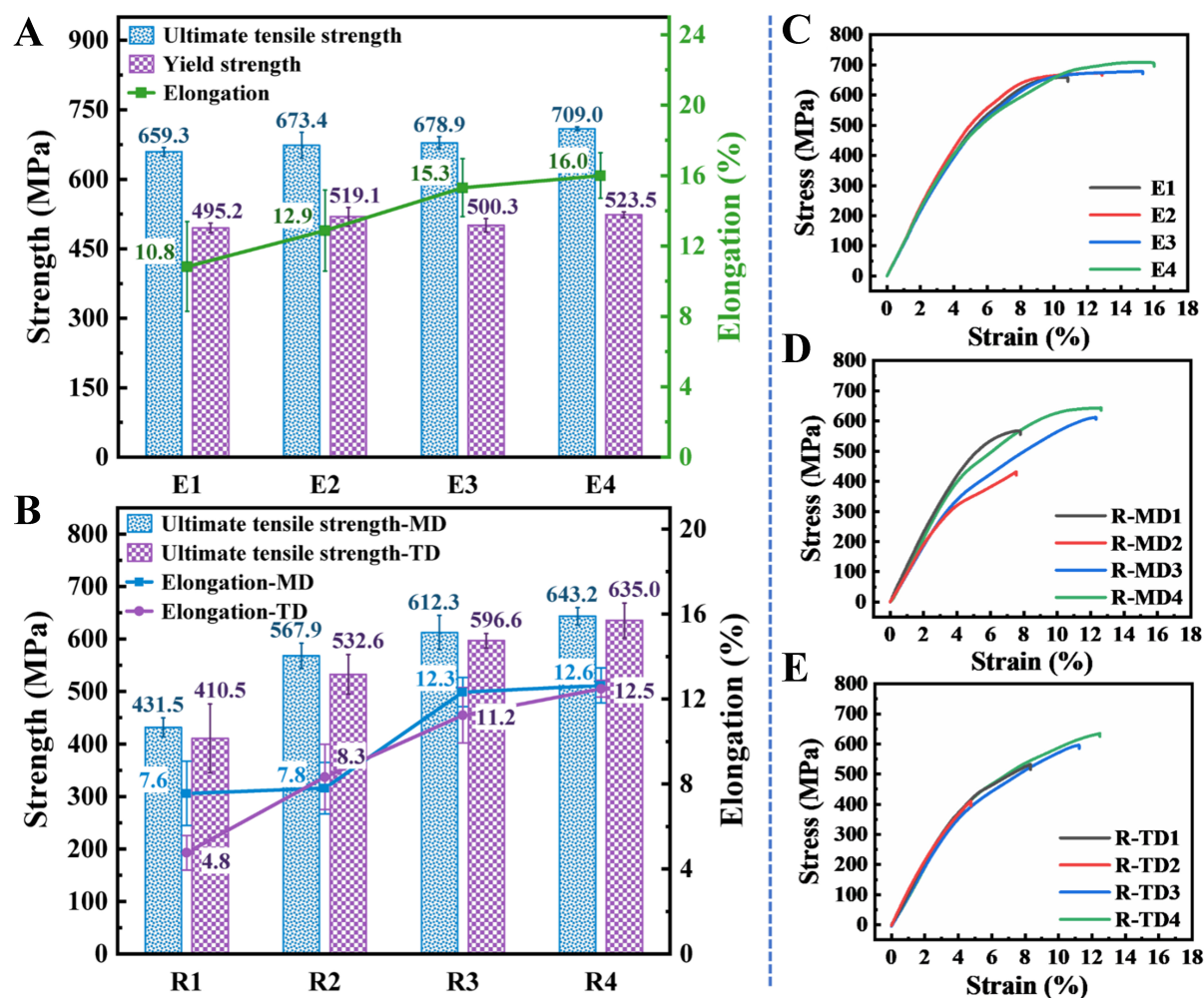


**Figure 13.** (A-D) Microstructure of as-cast samples without solid solution and ageing treatment; (E-H) and (I-L) Microstructure of extruded and rolled alloys after solid solution treatment, respectively. (A, E, and I) A1, (B, F, and J) A2, (C, G, and K) A3, (D, H, and L) A4.

agreement with the actual test results, thus verifying the validity of the model.

The strength of the alloys from E1 to E4 increases progressively, and the strength and EL of R1 to R4 also show an increasing trend. Comparing the two deformation processes, the mechanical properties obtained by the extrusion process (UTS of 650 to 710 MPa, EL of 10.8% to 16%) are significantly superior to those obtained by the rolling process (UTS of 410 to 650 MPa, EL of 5% to 13%). The superior mechanical properties obtained by extrusion may be due to the larger deformation volume compared to the rolling

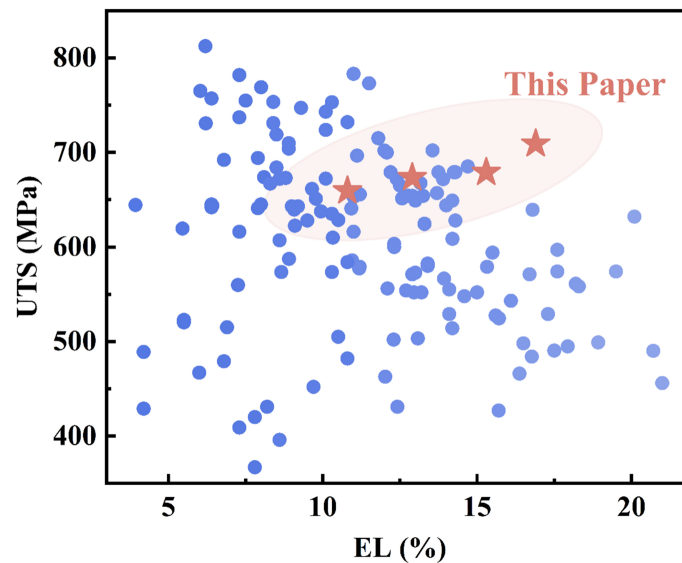




**Figure 14.** Mechanical properties of four alloys obtained by extrusion and rolling process. (A and B) stress-strain curve; (C-E) strength-EL plot. EL: Elongation.

process. This leads to a significant increase in dislocation density during the deformation process, resulting in the formation of ledges that hinder the movement of dislocations. As deformation proceeds, dislocation entangling makes dislocation slip more difficult, thereby greatly enhancing the strength. The four groups of alloy samples R1-R4 obtained by the rolling process, the machine direction (MD) strength was 431.5, 567.9, 612.3, 643.3 MPa, and the EL was 7.6%, 7.8%, 12.3%, and 12.6%, respectively. The transverse direction (TD) strengths were 410.5, 532.6, 596.6, and 635.0 MPa and ELs were 4.8%, 8.3%, 11.2%, and 12.5%, respectively. In the rolling process, the mechanical properties in the MD are better than those in the TD as this phenomenon is mainly related to the deformation method. During rolling, the deformation of the material along the rolling direction (i.e., MD) is usually larger and usually provides better mechanical properties. In contrast, the deformation in the TD is relatively small, and the grain shape and distribution are not as uniform as in the longitudinal direction, which leads to differences in the structure and weave, thus affecting the mechanical properties. In addition, the spreading ratio (TD) during rolling is usually smaller than the EL (MD), which partly explains why the MD properties are usually better than those in the TD<sup>[40]</sup>.

The mechanical properties of the alloy samples prepared from the extrusion experiments were compared with the mechanical properties of the alloys designed by ML [Supplementary Table 2]. A1 to A4 followed



**Figure 15.** The properties of the alloy designed in this paper are compared with the data from the database.

the same trend in strength change from E1 to E4 after extrusion, with the strength progressively increasing, which is basically the same as that of the alloy's mechanical properties. However, the experimental values are slightly lower than the predicted values and the EL after heat treatment is higher than the predicted design results; these cases may be due to the significant effect of the alloy deformation heat treatment process on the properties.

In this study, E4 alloys with a tensile strength of  $709 \pm 4$  MPa were obtained by extrusion process and reached an EL of  $16\% \pm 1\%$ . As shown in Figure 15, the designed alloy is compared with the aluminium alloys in the database established in this study. The E4 alloy shows excellent comprehensive mechanical properties. For detailed information on alloy composition and preparation process, refer to the detailed content in the data set in the Supplementary Materials. All the comparison alloys were prepared using conventional process conditions consistent with this study, and the mechanical property values were obtained at room temperature, following standard test methods widely recognised in the field. The same numerical units were used for all performance indicators to ensure the validity and reliability of the comparisons. The results show that the overall mechanical properties of Al-Zn-Mg-Cu alloys are significantly improved.

## CONCLUSIONS

In this study, a multi-algorithm integrated ML model was introduced to accelerate the screening of Al-Zn-Mg-Cu alloy components, and an alloy component design strategy was constructed. Four sets of optimal Al-Zn-Mg-Cu alloy compositions were designed. The main conclusions are as follows.

- (1) A dataset of compositions and properties of aluminium alloys was established. Based on this, the stacking multi-algorithm integrated model with the  $R^2$  of 0.96 was finally constructed for the screening of Al-Zn-Mg-Cu alloys.
- (2) Adopting the alloy design strategy of reverse design combined with forward prediction. Through this approach, four sets of high-strength, high-toughness compositions of Al-Zn-Mg-Cu alloys were successfully

designed.

(3) Experimentally prepared four sets of alloys and compared the two deformation processes, the mechanical properties obtained by the extrusion process (UTS of 650–710 MPa, EL of 10.8%–16%) are much better than those obtained by the rolling process (UTS of 410–650 MPa, EL of 5%–13%). Specifically, the E4 alloy exhibited a UTS of  $709 \pm 4$  MPa and an EL of  $16\% \pm 1\%$ , which represents a significant enhancement in comprehensive performance.

## DECLARATIONS

### Authors' contributions

Investigation, writing - original draft: Yuan Y

Methodology, visualisation, writing - review and editing: Sui Y, Jiang A

Data curation, methodology: Li P, Quan M

Conceptualisation, supervision: Zhou H

Conceptualisation, supervision, funding acquisition: Jiang A

### Availability of data and materials

The raw data were listed in the [Supplementary Materials](#). Further data used to support the findings of this study are available from the corresponding author upon reasonable request.

### Financial support and sponsorship

This work was supported by the Science and Technology Major Project of Yunnan Province (Grant No. 202202AG050011), National Natural Science Foundation of China (Grant No. U1902220), Yunnan Fundamental Research Projects (Grant No. 202201AT070099) and Project Funds of “Xingdian Talent Support Program”, Science and technology innovation project of Transportation Department of Yunnan Province (Grant No. Yun Jiao Ke Jiao Bian [2022] 123), and Science and Technology Innovation Plan Project of Yunnan Trading Group (YCIC-YF-2022-20).

### Conflicts of interest

All authors declared that there are no conflicts of interest.

### Ethical approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Copyright

© The Author(s) 2024.

## REFERENCES

1. Li S, Yue X, Li Q, et al. Development and applications of aluminum alloys for aerospace industry. *J Mater Res Technol* 2023;27:944–83. [DOI](#)
2. Starke E, Staley J. Application of modern aluminum alloys to aircraft. *Prog Aerosp Sci* 1996;32:131–72. [DOI](#)
3. Ni R, Boehlert CJ, Zeng Y, et al. Automated analysis framework of strain partitioning and deformation mechanisms via multimodal fusion and computer vision. *Int J Plasticity* 2024;182:104119. [DOI](#)
4. Yang H, Tian S, Gao T, et al. High-temperature mechanical properties of 2024 Al matrix nanocomposite reinforced by TiC network architecture. *Mat Sci Eng A* 2019;763:138121. [DOI](#)
5. Dursun T, Soutis C. Recent developments in advanced aircraft aluminium alloys. *Mater Design* 2014;56:862–71. [DOI](#)



6. Dang B, Zhang X, Chen YZ, Chen CX, Wang HT, Liu F. Breaking through the strength-ductility trade-off dilemma in an Al-Si-based casting alloy. *Sci Rep* 2016;6:30874. DOI PubMed PMC
7. Kim DW, Lee S, Sohn SS. Overcoming strength-ductility trade-off via subzero martensitic transformation in medium-Mn lightweight steel. *Scripta Mater* 2022;210:114477. DOI
8. Li Y, Wang Y, Lu B, et al. Effect of Cu content and Zn/Mg ratio on microstructure and mechanical properties of Al–Zn–Mg–Cu alloys. *J Mater Res Technol* 2022;19:3451-60. DOI
9. Zhang Z, Li Y, Li H, Zhang D, Zhang J. Effect of high Cu concentration on the mechanical property and precipitation behavior of Al–Mg–Zn-(Cu) crossover alloys. *J Mater Res Technol* 2022;20:4585-96. DOI
10. Tan P, Qin J, Quan X, Yi D, Wang B. Co-strengthening of the multi-phase precipitation in high-strength and toughness cast Al–Cu–Zn–Mg alloy via changing Zn/Mg ratios. *Mat Sci Eng A* 2023;873:145024. DOI
11. Liu C, Teng G, Ma Z, Wei L, Zhang B, Chen Y. Effects of Sc and Zr microalloying on the microstructure and mechanical properties of high Cu content 7xxx Al alloy. *Int J Miner Metall Mater* 2019;26:1559-69. DOI
12. Li W, Pan Q, Zou L, Liang W, He Y, Liu J. Effects of minor Sc on the microstructure and mechanical properties of Al–Zn–Mg–Cu–Zr based alloys. *Rare Metals* 2009;28:102-6. DOI
13. Wang Y, Cao L, Wu X, Lin X, Yao T, Peng L. Multi-alloying effect of Ti, Mn, Cr, Zr, Er on the cast Al–Zn–Mg–Cu alloys. *Mater Charact* 2023;201:112984. DOI
14. Wang Y, Wu R, Turakhodjaev N, Liu M. Microstructural evolution, precipitation behavior and mechanical properties of a novel Al–Zn–Mg–Cu–Li–Sc–Zr alloy. *J Mater Res* 2021;36:740-50. DOI
15. Li C, Zhou Q, Han M, Sha S, Luo Y, Xu X. Effect of rare earth element Er on the microstructure and properties of highly alloyed Al–Zn–Mg–Cu–Zr–Ti alloy. *J Alloys Compd* 2023;956:170248. DOI
16. Wu Y, Li C, Froes FH, Alvarez A. Microalloying of Sc, Ni, and Ce in an advanced Al–Zn–Mg–Cu alloy. *Metall Mater Trans A* 1999;30:1017-24. DOI
17. Li J, Zhang Y, Li M, Hu Y, Zeng Q, Zhang P. Effect of combined addition of Zr, Ti and Y on microstructure and tensile properties of an Al–Zn–Mg–Cu alloy. *Mater Design* 2022;223:111129. DOI
18. Chen Z, Mo Y, Nie Z. Effect of Zn content on the microstructure and properties of super-high strength Al–Zn–Mg–Cu alloys. *Metall Mater Trans A* 2013;44:3910-20. DOI
19. Huang R, Li M, Yang H, et al. Effects of Mg contents on microstructures and second phases of as-cast Al–Zn–Mg–Cu alloys. *J Mater Res Technol* 2022;21:2105-17. DOI
20. Berg LK, Gjønnes J, Hansen V, et al. GP-zones in Al–Zn–Mg alloys and their role in artificial aging. *Acta Mater* 2001;49:3443-51. DOI
21. Yang X, Chen J, Liu J, Qin F, Xie J, Wu C. A high-strength AlZnMg alloy hardened by the T-phase precipitates. *J Alloys Compd* 2014;610:69-73. DOI
22. Cheng LM, Poole WJ, Embury JD, Lloyd DJ. The influence of precipitation on the work-hardening behavior of the aluminum alloys AA6111 and AA7030. *Metall Mater Trans A* 2003;34:2473-81. DOI
23. Yin H, Wen K, Li Z, et al. Effect of Zn/Mg ratio on the microstructure and mechanical properties of as-cast Al–Zn–Mg–Cu alloys and the phase transformation during homogenization. *J Mater Res Technol* 2023;26:3646-60. DOI
24. Dong P, Chen S, Chen K. Effects of Cu content on microstructure and properties of super-high-strength Al–9.3Zn–2.4Mg–xCu–Zr alloy. *J Alloys Compd* 2019;788:329-37. DOI
25. Fu H, Zhang H, Wang C, Yong W, Xie J. Recent progress in the machine learning-assisted rational design of alloys. *Int J Miner Metall Mater* 2022;29:635-44. DOI
26. Xie J, Su Y, Zhang D, Feng Q. A vision of materials genome engineering in China. *Engineering* 2022;10:10-2. DOI
27. Chen Y, Wang S, Xiong J, et al. Identifying facile material descriptors for Charpy impact toughness in low-alloy steel via machine learning. *J Mater Sci Technol* 2023;132:213-22. DOI
28. Li J, Zhang Y, Cao X, et al. Accelerated discovery of high-strength aluminum alloys by machine learning. *Commun Mater* 2020;1:74. DOI
29. Ghorbani M, Boley M, Nakashima P, Birbilis N. A machine learning approach for accelerated design of magnesium alloys. Part A: alloy data and property space. *J Magnes Alloy* 2023;11:3620-33. DOI
30. Chaudry U, Hamad K, Abuhmed T. Machine learning-aided design of aluminum alloys with high performance. *Mater Today Commun* 2021;26:101897. DOI
31. Wang C, Fu H, Jiang L, Xue D, Xie J. A property-oriented design strategy for high performance copper alloys via machine learning. *npj Comput Mater* 2019;5:227. DOI
32. Chen Y, Tian Y, Zhou Y, et al. Machine learning assisted multi-objective optimization for materials processing parameters: a case study in Mg alloy. *J Alloys Compd* 2020;844:156159. DOI
33. Mi X, Tian L, Tang A, et al. A reverse design model for high-performance and low-cost magnesium alloys by machine learning. *Comp Mater Sci* 2022;201:110881. DOI
34. Jiang L, Wang C, Fu H, Shen J, Zhang Z, Xie J. Discovery of aluminum alloys with ultra-strength and high-toughness via a property-oriented design strategy. *J Mater Sci Technol* 2022;98:33-43. DOI
35. Jiang Y, Li Y, Liu F. Microalloying-modulated strength-ductility trade-offs in as-cast Al–Mg–Si–Cu alloys. *Mat Sci Eng A* 2022;855:143897. DOI

36. Marlaud T, Deschamps A, Bley F, Lefebvre W, Baroux B. Evolution of precipitate microstructures during the retrogression and re-ageing heat treatment of an Al–Zn–Mg–Cu alloy. *Acta Mater* 2010;58:4814-26. [DOI](#)
37. Liddicoat PV, Liao XZ, Zhao Y, et al. Nanostructural hierarchy increases the strength of aluminium alloys. *Nat Commun* 2010;1:63. [DOI](#) [PubMed](#)
38. Sui Y, Wang Q, Wang G, Liu T. Effects of Sr content on the microstructure and mechanical properties of cast Al–12Si–4Cu–2Ni–0.8Mg alloys. *J Alloys Compd* 2015;622:572-9. [DOI](#)
39. Qu Z, Zhang Z, Yan J, et al. Examining the effect of the aging state on strength and plasticity of wrought aluminum alloys. *J Mater Sci Technol* 2022;122:54-67. [DOI](#)
40. Xiang H, Xu C, Zhan T, Guo P, Li L. Fabrication of high strength-ductility aluminum alloy heterogeneous plates using additive manufacturing and hot rolling process. *J Mater Process Tech* 2024;329:118451. [DOI](#)