npj | Computational Materials

www.nature.com/npjcompumats

ARTICLE OPEN

# A property-oriented design strategy for high performance copper alloys via machine learning

Changsheng Wang[1], Huadong Fu[1], Lei Jiang[1], Dezhen Xue[2] and Jianxin Xie[1]

Traditional strategies for designing new materials with targeted property including methods such as trial and error, and experiences of domain experts, are time and cost consuming. In the present study, we propose a machine learning design system involving three features of machine learning modeling, compositional design and property prediction, which can accelerate the discovery of new materials. We demonstrate better efficiency of on a rapid compositional design of high-performance copper alloys with a targeted ultimate tensile strength of 600–950 MPa and an electrical conductivity of 50.0% international annealed copper standard. There exists a good consistency between the predicted and measured values for three alloys from literatures and two newly made alloys with designed compositions. Our results provide a new recipe to realize the property-oriented compositional design for high-performance complex alloys via machine learning.

*npj Computational Materials* (2019)5:87 ; https://doi.org/10.1038/s41524-019-0227-7

## INTORDUCTION

High-performance copper alloys are fundamental to the lead frames of integrated circuits (ICs). The rapid development of IC technology towards high density, multifunction, miniaturization and low cost, desires the copper alloys having higher performances on the mechanical strength, the electrical conductivity, etc.[1,2] For example, the traditional copper alloys, including Cu–Fe–P, Cu–Ni–Si and Cu–Cr–Zr alloys, are hardly be used in the next generation of very-large-scale integration (VLSI) ICs, which requires a ultimate tensile strength (UTS) over 800 MPa and an electrical conductivity (EC) over 50.0% International Annealed Copper Standard (IACS).[3–5]

To improve the mechanical and electrical properties of copper alloys, one or more alloying elements, such as Ti, Co, P, Mg, Cr, Zr, Be, and Fe, can be introduced. Many efforts have been devoted to this field and showed that the alloying elements should have little effect on the EC and possesses a large solid solubility change from high temperature to room temperature.[6–10] However, there is a lack of a model that quantitatively describes the relationship between alloy composition and performance. As a result, the compositional design of high-performance copper alloy mainly relies on *trial and error*, or intuitions. Especially, an object-oriented inverse design that recommends the compositions rapidly and accurately for desired properties is needed.[11–15]

Data mining or machine learning builds inference models that learn the relationship among the composition, processing conditions, microstructures, and properties of materials based on materials databases. This allows a "pre-design" strategy which designs materials before experiments, which is in contrast to the conventional "post-analysis" strategy.[16–20] Recently, neural network (NN), which is able to capture the highly complex non-linear input/output relationships, has been applied in materials science to build up the composition–processing–performance relationships and to directly predict the properties of alloys.[21,22] Reddy et al.[23] established an inference model from compositions and heat treatment conditions to mechanical properties of the low alloy steel by combining the back-propagation (BP) NN and genetic algorithm (GA). Their model successfully learns the influence of compositions and heat treatment conditions on the performance of the steel. Ozerdem et al.[24] built a multi-layer BP NN model to predict the yield strength, UTS and elongation of the Cu–Sn–Pb–Zn–Ni alloy. These NN models with inputs of compositions and processing conditions can estimate the properties of alloys. Such forward models from composition to property is helpful to screen or down-select the potential good candidates. However, more attractive thing is an inverse design model that recommends compositions from a targeted property, i.e., a property to composition predictive model. It allows a fast locating of property optima in the composition and processing search space, whereas is a tough problem due to the highly non-linear relationship among composition, processing, and performance.

In the present study, we propose a machine learning design system (MLDS) to realize property-oriented compositional design for high-performance complex alloys. Two kinds of BP NN models are built to learn the relationship between the materials properties including mechanical strength and electrical conductivity and the compositions of copper alloys based on a database with hundreds of samples. One of the models predicts properties of alloys from their compositions (i.e., composition → performance, C2P) and the other model predicts the compositions of alloys according to the targeted properties (i.e., performance → composition, P2C). For a given property requirement, the P2C model is utilized to screen the composition space, and then the C2P model is employed to efficiently recommend the targeted composition within a certain confidential interval. We demonstrate our MLDS by rapidly design of several high-performance copper alloys with an ultimate tensile

[1]Beijing Advanced Innovation Center for Materials Genome Engineering, Institute for Advanced Materials and Technology, University of Science and Technology Beijing, Beijing 100083, China and [2]State Key Laboratory for Mechanical Behavior of Materials, Xi'an Jiaotong University, Xi'an 710049, China
Correspondence: Jianxin Xie (jxxie@mater.ustb.edu.cn)
These authors contributed equally: Changsheng Wang, Huadong Fu

Received: 31 January 2019 Accepted: 7 August 2019
Published online: 27 August 2019

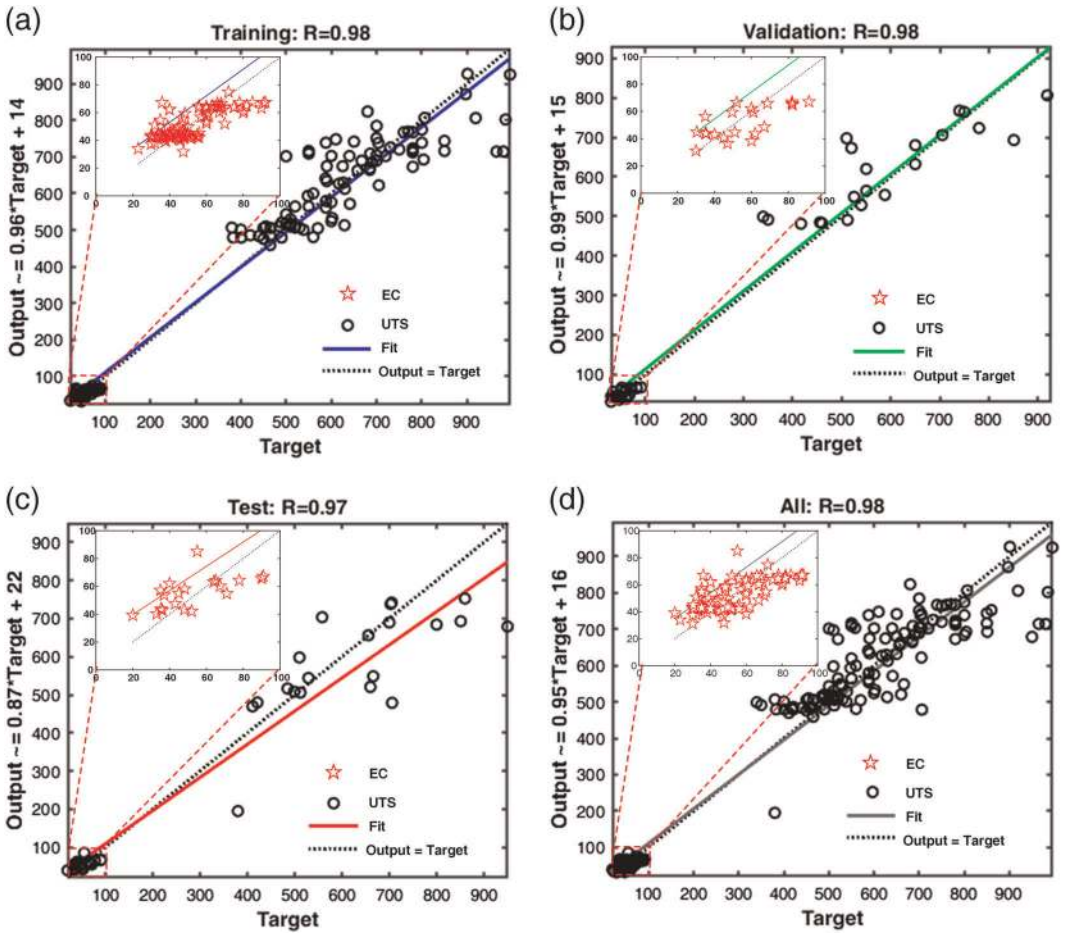


**Fig. 1** The predicted UTS and EC values from the composition → property (C2P) model as a function of the measured UTS and EC values for **a** the training set, **b** the validation set, **c** the test set, and **d** all data. (The dotted line represents the output data of the model in exactly the same as the target data in the data set; the solid line represents the regression results between the output data and the target data.)

strength (UTS) of 600–950 MPa and an electrical conductivity (EC) of 50.0% IACS. The present strategy allows a rapid and accurate compositional design of high-performance, multi-component, and complex copper alloys.

## RESULTS AND DISCUSSION

### Construction and analysis of two BP NN models

The first BP NN (C2P) model is trained to predict properties of UTS and EC from compositions of copper alloys on a condition of keeping the processing unchanged. The input is compositions and the outputs are UTS and EC, which were shown in Fig. S1 and Fig. S2. More details of the C2P model are shown in the supplementary information.

Figure 1a–d plots the predicted UTS and EC values from the model as a function of the measured UTS and EC values for the training set, the validation set, the test set, and all data, respectively. Data points fall along the diagonal line (as shown by the dotted line in each panel), indicating that the predictions are quite consistent with the measured values. We further linearly fitted the predicted values with respect to the measured values, as shown by the solid line. Closer the solid line to the diagonal dotted line is, better the performance of the model is. It can be seen that our model performs very good on separate data sets and on all data.

The inverse design of the composition, processing, and microstructure of alloys with targeted properties will greatly reduce the time and cost for developing new materials and is the ultimate goal of material science.[25] The greatest obstacle is the complex relationship among the composition, processing, structure, and performance. When the processing conditions are given constantly, the composition is the dominant variable that determines the microstructures and the properties of an alloy. Therefore, we build the BP NN model (P2C) with the input of UTS and EC and output of alloying elements except Cu, which were shown in Fig. S3 and Fig. S4. More details of the P2C model are shown in the supplementary information.

The scatter plots in Fig. 2 show that the predicted content values are very consistent with the measured values and fall along the dotted diagonal lines. Moreover, the linearly fitted solid curves are very close to the dotted diagonal lines. Thus, the P2C BP NN performs fairly well and can be utilized to recommend new compositions. The P2C model can simultaneously output a set of 10 elements' concentration values when it is running. For example, when we input UTS and EC values to the P2C model, the P2C model outputs 10 elements' concentration values at the same time. A set of 10 elements' concentration values is called an alloy composition design scheme.

We further validate the prediction accuracy of C2P model and P2C model by comparing the predictions with the experimental data which are not in our database. The performance on training data is also included in this section. Table S1 in the supplementary information lists the root mean square error (RMSE) of the C2P model and the P2C model trained with the same database, and

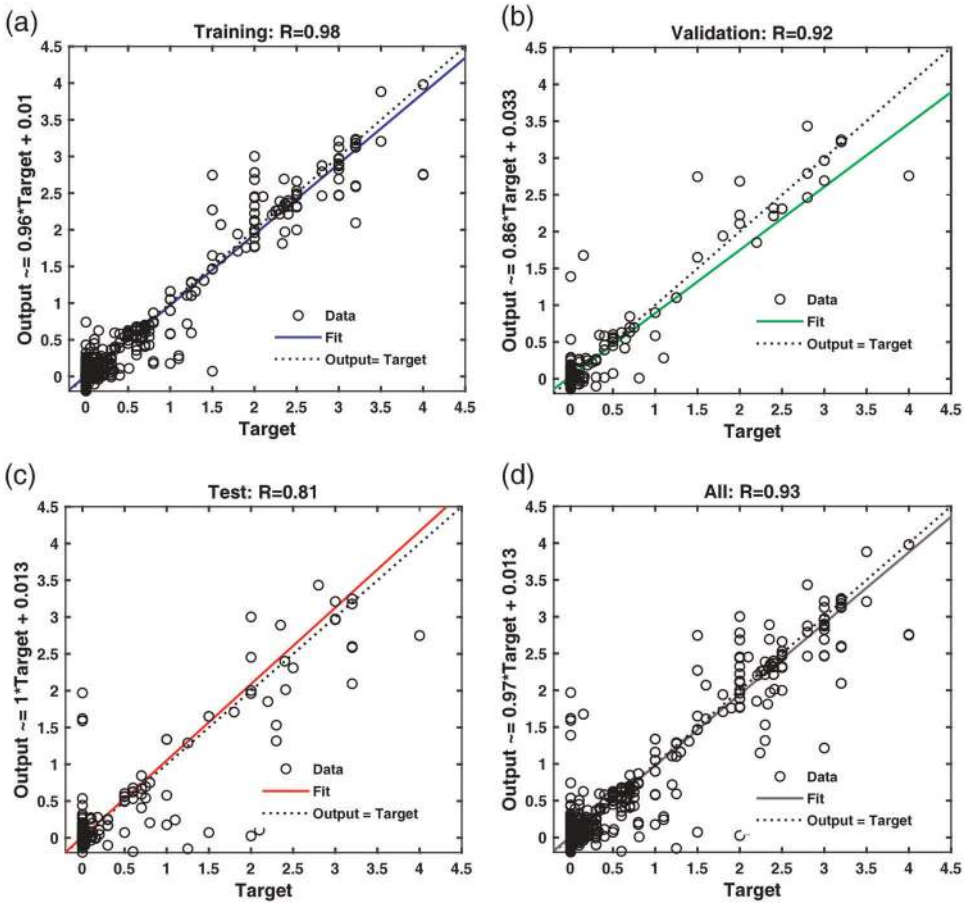


**Fig. 2** The predicted content values from the property → composition (P2C) model as a function of the measured content values for **a** the training set, **b** the validation set, **c** the test set, and **d** all data. (The dotted line represents the output data of the model in exactly the same as the target data in the data set; the solid line represents the regression results between the output data and the target data.)

the regression fit correlation data (R) between predicted performance and target performance under different data sets. The C2P network training error accuracy and regression fitting correlation coefficients on verification set and test set are higher than the P2C model. The above results show that both models perform well in terms of the training performance, and the C2P model is still better than the P2C model. The main reasons for this difference are discussed further below.

In order to validate the predictive performance of C2P model, we chose alloys that are not presented in our database. There are many factors affecting the properties of the copper alloys, especially the processing conditions. The alloy with the same composition could have different microstructures produced by different processing conditions, leading to a large deviation in the properties. We use the best properties of these alloys with the same composition after solution treatment, deformation, and aging treatment to validate our C2P model.

The predicted and experimental values for these alloys are compared in Table 1. Cu–2.93Ni–0.9Si–0.13Mg–0.53Zn alloy is a commercialized product and Cu–2.20Ni–0.42Si–0.08Mg–0.30Zn alloy is made and measured in our own lab and the rest four alloys[26–28] are reported very recently. These data are not included in our database. Table 1 shows that the deviation between predicted and experimental values are within 5.0%, indicating that the C2P model can accurately predict the properties of alloys.

In order to validate the accuracy of alloy composition designed by the P2C model, we set the targeted UTS as 500, 550, 600, 650, and 700 MPa, respectively, together with the targeted EC of 50.0% IACS. These properties are input into the P2C model and some alloy composition design schemes are recommended, as shown in Table 2. The elements in brackets were ≤0.01 wt%, which can be considered to be trivial.

Table 2 shows that with targeted UTS increasing from 550 to 700 MPa, the recommended alloys by P2C model change from Cu–Fe–P series alloy (1#) to Cu–Ni–Si series alloy (2–5#). For 2–5# alloys, the UTS highly correlates with the content of Ni; but there is no obvious change in other alloying elements as those <0.01 wt% are considered as the impurity elements. We found there exist four experimental alloys in literatures which are very close to the five designed alloys in Table 2. Their properties and compositions also listed in Table 3. The best UTS and EC values of Cu–2.35Fe–0.10Cr–0.03P[29] alloy after homogenization, hot rolling, solution treatment, cold rolling, and aging treatment are 495 MPa and 58.0% IACS, respectively. Its compositions are very similar with the 1# alloy in Table 2. The UTS is very close (the difference is only 5 MPa), and the EC is 8.0% IACS higher than that of 1# alloy. The contents of Ni, Si, and Mg of the second,[30] third[31], and forth[32] alloys in Table 3 are similar to those of 2#, 3#, and 4# in Table 2. The measured EC of the second alloy is 4.1% IACS lower than the target EC of 2# alloy, and the measured UTS is 120 MPa higher than that of 2# alloy (of 21.8%). The experimental EC of the third alloy is 6.0% IACS (of 12.0%) lower than of that of 3# and 4# alloys, and the experimental UTS is 30 MPa higher than that of 3# alloy and 20 MPa lower than that of the 4# alloy. The composition of Cu–4.00Ni–0.95Si–0.02Cr–0.02P[32] alloy in Table 3 is similar to 5# alloy in Table 2, but their properties are quite different. The

**Table 1.** Validation of prediction of alloy performance based on composition (C2P model)

| Alloy composition/wt% | Predictive value | | Experimental value | |
|---|---|---|---|---|
| | UTS/MPa | EC/%IACS | UTS/MPa | EC/%IACS |
| Cu–2.93Ni–0.9Si–0.13Mg–0.53Zn | 646 | 37.6 | 627 ± 11 | 37.5 ± 0.4 |
| Cu–2.20Ni–0.42Si–0.08Mg–0.30Zn | 702 | 47.6 | 720 ± 12 | 49.5 ± 0.5 |
| Cu–2.69Ni–1.14Si–0.45Cr[26] | 780 | 42.8 | 757 ± 13 | 44.0 ± 1.0 |
| Cu–2.11Ni–0.54Si–0.21Cr–0.16Zr[27] | 654 | 47.4 | 625 | 48.8 |
| Cu–0.42Cr[28] | 479 | 83.2 | 457 | 85.1 |
| Cu–0.28Cr–0.15Mg[28] | 536 | 82.6 | 515 | 80.8 |

**Table 2.** Recommended compositions according to the targeted properties of UTS and EC based on the P2C model

| No. | UTS/MPa | EC/%IACS | Predicted alloy composition/wt% |
|---|---|---|---|
| 1# | 500 | 50.0 | Cu–2.66Fe–0.02P–0.1Cr(–0.01Zn) |
| 2# | 550 | 50.0 | Cu–1.98Ni–0.45Si(–0.001Mg–0.001Zn) |
| 3# | 600 | 50.0 | Cu–2.00Ni–0.60Si(–0.01Mg–0.01Sn–0.01Fe) |
| 4# | 650 | 50.0 | Cu–2.20Ni–0.48Si–0.1Mg(–0.01Fe–0.001 P) |
| 5# | 700 | 50.0 | Cu–3.50Ni–0.85Si–0.03Cr (–0.01Zn–0.01P–0.01Sn) |

**Table 3.** Alloys with similar compositions with those listed in Table 2 and their reported properties in literatures

| Alloy composition | UTS/MPa | EC/%IACS |
|---|---|---|
| Cu–2.35Fe–0.03P–0.10Cr–0.03 P[29] | 495 | 58.0 |
| Cu–2.00Ni–0.50Si[30] | 670 | 45.9 |
| Cu–2.00Ni–0.50Si–0.10Mg[31] | 630 | 44.0 |
| Cu–4.00Ni–0.95Si–0.02Cr–0.02P[32] | 830 | 35.0 |

measured EC of this alloy is 15.0% IACS lower than 5# alloy, and the measured UTS is 130 MPa higher than 5# alloy (of 18.6%).

It can be seen that two BP NN models (C2P model and P2C model) trained by the same data set deviate in training performance and predictive ability. The C2P model can obtain satisfactory performance prediction results, and the composition design results obtained by the P2C model is less reliable. In machine learning, when the number of training samples is constant, the prediction ability of the model decreases with the increase of the prediction dimension (the number of output variables).[33] We anticipate that the main reason for such a difference relies on the different topologies of the two models. The C2P model uses more than ten independent variables to fit two dependent variables, which is a "dimension-reduction fitting". The training of model is easy to converge. The P2C model fit more than ten dependent variables with only two independent variables, which is a "dimension-increase fitting". The convergence of model is poor and it is prone to be over-fitting. In most cases, high-performance alloys are multi-components. As a result, using P2C model to inversely design the composition does easily encounter the "dimension-increase fitting" mentioned above, and consequently it is difficult to obtain satisfactory compositions. This could be an important reason for the slow progress in inverse design of alloy compositions so far. In order to solve this problem, it is necessary to develop a robust compositional design approach to rapidly and accurately design a reasonable alloy with targeted properties.

## Construction and application of a property-oriented MLDS

To this end, we propose the following machine learning design system (MLDS) and its flow chart is shown in Fig. 3. The basic idea is for high reliability of C2P model and low reliability of the P2C model, we use C2P model to evaluate the alloy composition design schemes provided by P2C model, and establish the criterion of whether the composition design is reasonable (i.e., the relative error between target performance and predictive performance). Through repeated *trial and error* of MLDS system, a more reasonable alloy composition design scheme can be screened automatically and quickly. The specific methods are as follows:

First, two models of C2P and P2C are established by BP NN. Second, the initial alloy composition design scheme is performed by P2C model based on targeted properties. Third, the previous schemes are input into the C2P model with higher reliability to get a more accurate prediction. Fourth, the predicted properties are compared with the targets and an error is obtained. According to the error, either a selection of composition design scheme or a retraining of the model will be conducted. If all the errors between predicted and targeted values exceed a preset threshold, the P2C model will be retrained until a reasonable alloy composition design scheme is screened out. The MLDS is divided into three subsystems, including the model training, the compositional design and the property prediction. The three subsystems are interrelated and form a closed system.

To compare the performance of the MLDS and solely the P2C model on compositional design, we set targeted properties of UTS to 750 MPa and EC to 50.0% IACS. In total, 20 possible alloys were designed by MLDS with error threshold of 10% and P2C model and are listed in Table S2.

The concentration of main components (Ni, Si, Zn, and Cr) of the alloys designed are plotted in Fig. 4. The fluctuations of Ni, Si, Zn, and Cr of the alloys designed by P2C model are significant, while the fluctuations of the alloys designed by MLDS are small and stable. As shown in Table S2 the average concentrations of Ni, Si, Zn, and Cr of the alloys designed by MLDS are 2.29 wt%, 0.51 wt%, 0.25 wt%, and 0.06 wt%, respectively. While those of the alloys designed by P2C model are 2.46 wt%, 0.55 wt%, 0.30 wt%, and 0.12 wt%, respectively, which are 0.17 wt%, 0.04 wt%, 0.05 wt%, and 0.06 wt% higher, compared with alloys designed by MLDS. In addition, in terms of the standard deviation, the maximum value and minimum value of the predictions (i.e., the fluctuation range of the composition design), the MLDS performs much better than the P2C model. The above results show that MLDS has higher efficiency and reliability in solving the problem of compositional design for targeted properties.

We now set the targeted EC of 50.0% IACS and UTS of 600–950 MPa and conduct the MLDS with error threshold of 10.0%. The alloy composition design schemes are shown in Table 4.[34–36] With the increase of the targeted value of UTS, the concentrations of Ni, Si, and Cr elements become larger. In fact, the Ni, Si, and Cr elements are well-known strengthening

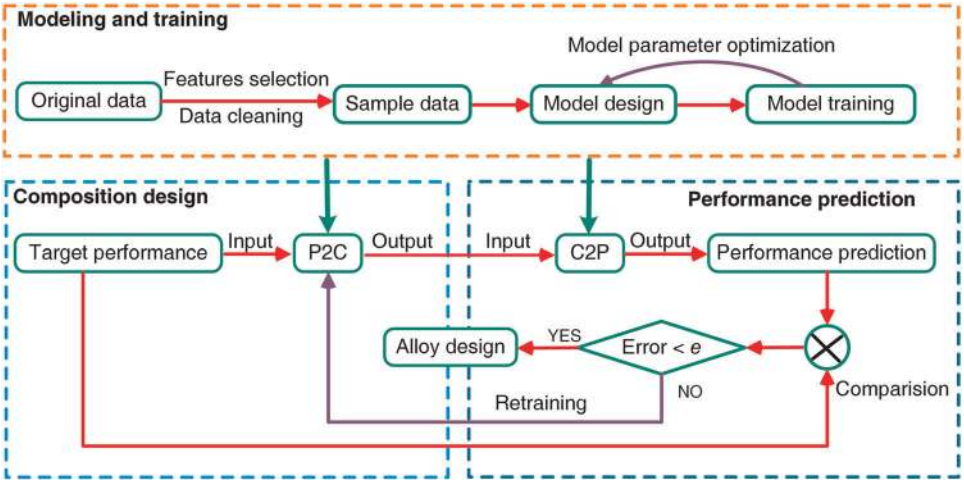


**Fig. 3** Flow chart of the machine learning design system (MLDS) for rapid and accurate compositional design

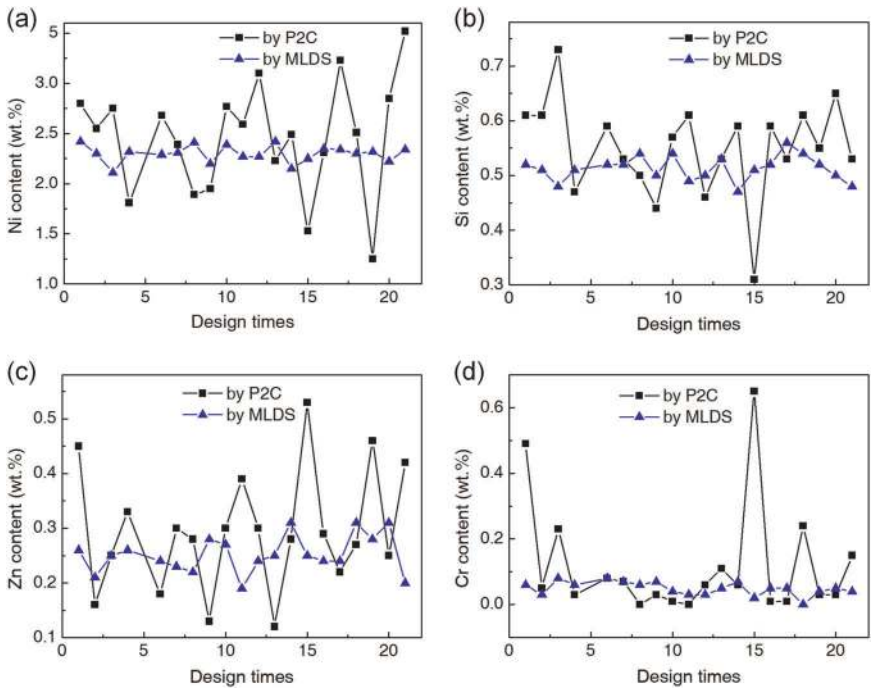


**Fig. 4** Main components, **a** Ni, **b** Si, **c** Zn, and **d** Cr, of alloys designed by P2C model and MLDS with the targeted properties as UTS of 750 MPa and EC of 50.0% IACS

**Table 4.** Design results and verification of composition of high-strength and high-electrical-conductivity multi-element complex copper alloys

| No. | Targeted properties | | Composition designed by MLDS/wt% | | | | | | | Validation of reference or experimental results | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | UTS/MPa | EC/IACS% | Ni | Si | Mg | Zn | Sn | Cr | P | Alloy | UTS/MPa | EC/IACS% |
| 1# | 600 | 50.0 | 1.47 | 0.39 | – | 0.18 | 0.02 | – | – | Cu–1.37Ni–0.28Si–0.02Mg–0.04Zn–0.04Sn[34] | 607 | 53.0 |
| 2# | 650 | 50.0 | 1.85 | 0.43 | – | 0.18 | – | 0.04 | 0.02 | – | – | – |
| 3# | 700 | 50.0 | 1.93 | 0.54 | 0.04 | 0.16 | 0.09 | – | 0.02 | Cu–2.00Ni–0.50Si–0.30Cr[35] | 700 | 49.7 |
| 4# | 750 | 50.0 | 2.32 | 0.51 | 0.04 | 0.26 | – | 0.06 | 0.02 | Cu–2.20Ni–0.42Si–0.08Mg–0.30Zn (Ref. Table 1) | 720 ± 12 | 49.5 ± 0.5 |
| 5# | 800 | 50.0 | 3.00 | 0.60 | – | 0.16 | – | 0.15 | 0.03 | Cu–3.00Ni–0.60Si–0.16Zn–0.15Cr–0.03P (Ref. Fig. 5) | 775 ± 10 | 48.0 ± 0.5 |
| 6# | 850 | 50.0 | 3.55 | 0.77 | 0.09 | 0.22 | – | 0.08 | 0.09 | – | – | – |
| 7# | 900 | 50.0 | 4.25 | 0.91 | 0.07 | 0.07 | – | 0.16 | 0.11 | Cu–4.00Ni–1.00Si–0.10Mg[36] | 951 | 44.0 |
| 8# | 950 | 50.0 | 4.61 | 0.92 | 0.11 | 0.08 | 0.13 | 0.20 | 0.13 | – | – | – |

elements for copper alloys, and can form strengthening phases such as $Ni_2Si$ and $Cr_3Si$, leading to a significant increase of the UTS.[2,26] The main role of trace Zn and Sn elements is to increase the energy of the stacking faults of the matrix, which improves the UTS, but has a negative influence on the EC.[1] The main role of trace Mg and P elements is to purify the melt and increase the fluidity of the melt during melting and casting.[32,35] It is also shown that the addition of P can promote the precipitation of the strengthening phases and Cr particles,[34,35] and thus can enhance the strengthening effect of Ni, Si, and Cr. It also benefits the EC of the alloy. Therefore, in the 5#–8# alloys, with the concentration increase of the strengthening elements such as Ni, Si, and Cr, the P, the strength is significantly increased as well. In summary, in Table 4, Ni, Si, and Cr are the main strengthening elements, while Mg, P, Zn, and Sn are auxiliary elements, which are used to balance trade-off between the mechanical properties and electrical conductivity of the alloy. In Table 4, the Fe element is not present in the design composition of the 1#–8# alloy because the strengthening effect of the Fe is not as remarkable as Ni, Si, and Cr. Although adding more Fe can improve the UTS of the copper alloy, excessive addition of Fe will have a negative influence on the EC of the alloy. In fact, the current Cu–Fe–P alloy series are of low UTS and high EC.

As shown in Table 4, the content of Ni, Si, and Cr, as the major strengthening elements, increases significantly with the increase of the target UTS in copper alloys designed by MLDS model, and the mass ratio of Ni/Si remains between of 3.5 and 5.0. These results are consistent with the basic principles of copper alloy design based on metallurgy.[37,38] On the other side, when the content of strengthening elements such as Ni, Si, and Cr increases significantly, the EC of copper alloys will inevitably be adversely affected. To solve this contradiction, the scheme recommended by MLDS is to increase the concentration of P element, as the trace P element can purify the matrix and promote the precipitation of the second phase of Cu–Ni–Si alloys, which is beneficial to improving the EC of Cu–Ni–Si alloys.[39,40] Thus, the results obtained by using MLDS model for alloy composition design conform to the basic principles of metallurgy, and further confirm the feasibility of alloy composition design by MLDS method based on machine learning.

Three alloys with similar principal components collected in literatures are listed in Table 4. Their processing conditions are solution treatment, deformation, and aging. And they are absent in our initial database shown in Fig. 6. The UTS and EC of the three alloys in Table 4 are very close to the targeted properties we set for MLDS, demonstrating the reliability of MLDS and the rationality of the design results.

Specifically, the concentration of Ni and Si in Cu–1.37Ni–0.28Si–0.02Mg–0.04Zn–0.04Sn–0.03P[34] alloy is similar to that of Ni and Si in 1# alloy designed by MLDS. Its UTS is very close to that of 1# alloy, and its EC is 3.0% IACS higher than that of 1# alloy. The main reason is that the concentration of Zn in the design alloy is high and there is no P element which can improve the EC of the alloy. The concentration of Ni and Si in Cu–2.00Ni–0.50Si–0.30Cr[35] is very close to that of Ni and Si in 3# alloy. Although the 3# alloy does not contain Cr, the total content of Mg, Zn, Sn, and P elements is 0.31 wt%, which is almost equal to Cr (0.30 wt%) of the alloy in literature. Although 3# alloy does not contain significant strengthening effect of Cr, and contains more Mg, Zn, and Sn which also have solid solution strengthening effect, the UTS and the EC of the alloy are almost the same as the target properties of 3# alloy in Table 4. Trace P can promote the precipitation of δ-$Ni_2Si$, which is conducive to the EC of 3# alloy. For Cu–4.00Ni–1.00Si–0.10Mg[36] alloy, the Ni concentration is slightly lower than that of 7# alloy, and Si and Mg concentration is slightly higher than that of 7# alloy. As a result, the UTS is higher than the targeted UTS of 7# alloy, and the EC is lower than the targeted EC of 7# alloy. In addition, Cr and Si of 7# alloy can form a strengthening phase $Cr_3Si$, which is also beneficial to the EC. Micro-alloyed P also has a positive effect on EC, resulting in a large difference (the relative error is 12.0%) between the target EC and the measured EC of 7# alloy.

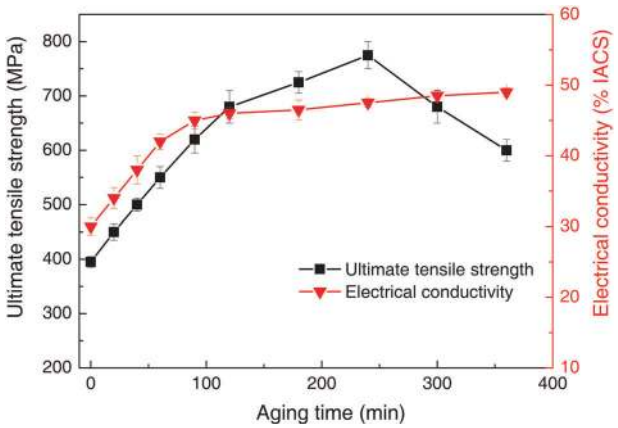


**Fig. 5** The ultimate tensile strength as a function of aging time at 450 °C of Cu–3.00Ni–0.60Si–0.16Zn–0.15Cr–0.03 P alloy

To further validate the rationality of the compositional design by MLDS, we prepared and measured two of the recommended alloys in our own lab, which are also listed in Table 4. The Cu–2.20Ni–0.42Si–0.08Mg–0.30Zn alloy has been used to evaluate the C2P model in Table 1. The concentration of its main components is close to that of the 4# alloy. The error between the measured properties (UTS of 720 MPa and EC of 49.5% IACS) and the targeted properties of 4# alloy (UTS of 750 MPa and EC of 50.0% IACS) is <4.0%. The second alloy of Cu–3.00Ni–0.60Si–0.16Zn–0.15Cr–0.03P was prepared according to the composition of 5# alloy in Table 4. The UTS and EC of the samples were measured and the UTS results as a function of aging time are shown in Fig. 5. The properties are optimized after aging for 240 min, which gives the UTS of 775 ± 10 MPa, and the EC of 48.0 ± 0.5% IACS. The difference between the measured and targeted properties is 25 MPa and 2.5% IACS (the relative errors of 3.1% and 4.0%), respectively. The measured properties according to the designed compositions are in good agreement with the target properties.

The MLDS shown in Fig.3 in the present study has high reliability in alloy composition design, and can satisfy the requirements of rapid and accurate compositional design for given property requirements (targeted properties).

In summary, we first proposed a machine learning design system that combines a property-composition BP NN model that designs the compositions and a composition-property BP NN model to efficiently screen the alloy composition design schemes. A rapid and more accurate design of alloy compositions is demonstrated. The MLDS provides a new method for accelerating the discovery of new materials.

Next, by using the MLDS, the inverse design of eight new copper alloys with UTS of 600–950 MPa and EC of 50.0% IACS is realized, which can provide good candidates for the research and development of high-performance copper alloys for large-scale integrated circuit lead frame and high-end connectors.

Finally, the rationality of the compositional design of the MLDS is validated by collecting three alloys from literatures and preparing two new alloys in our own laboratory according to the recommended compositions by the MLDS. Errors between measured properties and the target properties are <6%, indicating a satisfactory compositional design ability of the MLDS.

**Table 5.** The upper and lower limits of various elements and the range of the properties of UTS and EC in the training data

| Content of alloy elements (wt%) | | | | | | | | | | Performance | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Fe | P | Ni | Si | Mg | Zn | Sn | Cr | Zr | RE | UTS/MPa | EC/%IACS |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 340 | 22.0 |
| – | – | – | – | – | – | – | – | – | – | – | – |
| 2.50 | 0.32 | 4.00 | 1.80 | 0.30 | 2.50 | 1.25 | 0.90 | 0.30 | 0.81 | 1020 | 92.0 |

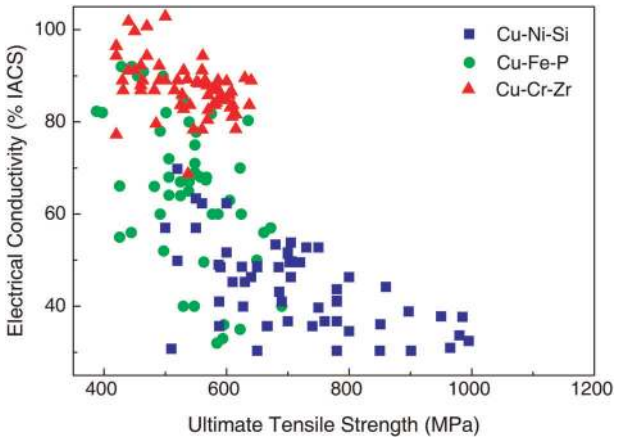


**Fig. 6** Ashby plot with respect to the electrical conductivity and ultimate tensile strength from all data in our database

## METHODS

### Data preparation

Developing copper alloys with high-strength and high-electrical-conductivity has been an attractive field for decades and has accumulated a large amount of literatures. We collected the data from these literatures of Cu–Fe–P, Cu–Ni–Si, and Cu–Cr–Zr based copper alloys, including the compositions and the corresponding materials properties of UTS and EC. To ensure that the recommended alloys can be synthesized experimentally in our own lab, we restrict the data from alloys with conventional strengthening methods, including solid-solution, precipitation, and deformation. Although some of alloy process parameters studied in the literature are not clear, they all adopt thermo-mechanical treatment process. Ignoring the different effects of process parameters, the default process is consistent. We abandon the data of alloys strengthened by unconventional methods, such as cryogenic treatment, equal channel angular pressing, high pressure torsion, and accumulative rolling bonding.[32,41–43] Moreover, data from alloys containing the toxic elements of Be and Cd, and the noble metal elements of Ag and Pt are also discarded.[44] Finally, a database containing around 300 well-labeled samples is established for the following machine learning process. Table 5 lists the upper and lower limits of various elements and the range of the properties of UTS and EC within our database. Figure 6 shows the Ashby plot of UTS and EC for all the data in our database and the empty area of upper-right is our design target.

### Modeling method

BP NN was proposed by Rumelhart and McCellando, which is a multi-layer feedforward network trained by error reverse propagation algorithm, and it is one of the most widely used NN models at present. BP NN uses the fastest descent method to learn the input and output data, and continuously adjusts the weights and thresholds of the network through the reverse propagation of errors, so that the error squared and minimum of the network.[45–47]

In principle, the BP NN is able to approximate any function and map highly non-linear relationships. The method of BP NN modeling is a statistical modeling method based on more data, and the result is a statistical law. This method does not require that all sample data (data set data) are generated under strictly comparable conditions, and the results obtained can not be called exact solutions in the strict sense. But when there is enough data, the accuracy of the results can be high enough, which is why big data's method can solve many problems that are difficult to solve by strict mathematical methods. The BP NN model includes an input layer, hidden layers, and an output layer.[48,49] The BP NN module of MATLAB was used to build two kinds of BP NN models of C2P and P2C.

The collected database is normalized to avoid possible influences from the large deviation in magnitude among the compositions and properties. The database is randomly divided into a training set, a verification set, and a test set according to a ratio of 60%, 20%, and 20%, respectively. The verification set is used to determine the best training times and prevent the over-fitting, and the test set is used to test generalization performance of the networks. It is noted that the results of the models trained at different times are slightly different, as the separation of the three sets are random.

### Experimental procedures

The preparation procedure of the experimental alloy is the same as the main procedure in the actual industrial production, as follows: about 99.99% pure cathodic electrolytic copper, Ni, Si, Zn, Cr, and Cu–14P were selected as raw materials, and Cu–3.00Ni–0.60Si–0.16Zn–0.15Cr–0.03P alloy was prepared by induction melting in ZG-25 medium frequency vacuum introduction furnace; the as-cast ingot was homogenized at 950 °C for 2 h in SGM-M30/12A resistance furnace, and then rapidly hot rolled at 850 °C from a thickness of 50 to 10 mm (80% reduction), followed by water quenching; the hot-rolled plate was then cold rolled with 90% reduction in thickness (1 mm) and then aged at 450 °C for 0–6 h. The UTS values of aged samples were tested by CMT6000 testing machine with a constant strain rate of 0.5 mm/min, and EC values were measured by Sigma2008B eddy current conduct-meter.

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ACKNOWLEDGEMENTS

This work was supported by the National Key Research and Development Program of China (No. 2016YFB0301300), and the National Natural Science Foundation of China (No. 51504023 and U1602271).

## AUTHOR CONTRIBUTIONS

C.S.W. and H.D.F. contributed equally to this work. C.S.W. and H.D.F. carried out the majority of the modeling and experimental work. J.X.X. designed modeling approaches and supervised the research. All authors interpreted the results and contributed to writing the paper.

## ADDITIONAL INFORMATION

**Supplementary Information** accompanies the paper on the *npj Computational Materials* website (https://doi.org/10.1038/s41524-019-0227-7).

**Competing interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

1. Gorsse, S., Ouvrard, B. & Gouné, M. Microstructural design of new high conductivity-high strength Cu-based alloy. *J. Alloy. Compd.* **633**, 42–47 (2015).
2. Yi, J. et al. Precipitation behavior of Cu-3.0Ni-0.72Si alloy. *Acta Mater.* **166**, 261–270 (2019).
3. Maki, K., Ito, Y., Matsunaga, H. & Mori, H. Solid-solution copper alloys with high strength and high electrical conductivity. *Scr. Mater.* **68**, 777–780 (2013).
4. Xu, S., Fu, H., Wang, Y. & Xie, J. Effect of Ag addition on the microstructure and mechanical properties of Cu-Cr alloy. *Mater. Sci. Eng. A* **726**, 208–214 (2018).
5. Mishnev, R., Shakhova, I., Belyakov, A. & Kaibyshev, R. Deformation microstructures, 0strengthening mechanisms, and electrical conductivity in a Cu-Cr-Zr alloy. *Mater. Sci. Eng. A* **629**, 29–40 (2015).
6. Dong, Q. et al. Microstructure and properties of Cu-2.3Fe-0.03P alloy during thermomechanical treatments. *Trans. Nonferrous Met. Soc. China* **25**, 1551–1558 (2015).
7. Li, D. et al. Minor-alloyed Cu-Ni-Si alloys with high hardness and electric conductivity designed by a cluster formula approach. *Prog. Nat. Sci.* **27**, 467–473 (2017).
8. Zhang, Y. et al. Dynamic recrystallization model of the Cu-Cr-Zr-Ag alloy under hot deformation. *J. Mater. Res.* **31**, 1–11 (2016).
9. Saravanan, K. et al. Studies on dynamic elastic and internal friction properties of Cu-Cr-Zr-Ti alloy between 25 and 650 °C. *J. Mater. Eng. Perform.* **24**, 4721–4727 (2015).
10. Lei, Q. et al. Hot working behavior of a super high strength Cu-Ni-Si alloy. *Mater. Des.* **51**, 1104–1109 (2013).
11. Sparks, T. D., Gaultois, M. W., Oliynyk, A., Brgoch, J. & Meredig, B. Data mining our way to the next generation of thermoelectrics. *Scr. Mater.* **111**, 10–15 (2016).
12. Ramprasad, R., Bratra, R., Pilania, G., Mannodi-Kanakkithodi, A. & Kim, C. Machine learning in materials informatics: recent applications and prospects. *npj Comput. Mater.* **3**, 54 (2017).
13. Verpoort, P. C., MacDonald, P. & Conduit, G. J. Materials data validation and imputation with an artificial neural network. *Comput. Mater. Sci.* **147**, 176–185 (2018).
14. Ward, L., Agrawal, A., Choudhary, A. & Wolverton, C. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Comput. Mater.* **2**, 16028 (2016).
15. Ward, L. et al. Matminer: an open source toolkit for materials data mining. *Comput. Mater. Sci.* **152**, 60–69 (2018).
16. Xue, D. et al. Accelerated search for materials with targeted properties by adaptive design. *Nat. Commun.* **7**, 11241 (2016).
17. Hu, X. et al. Two-way design of alloys for advanced ultra supercritical plants based on machine learning. *Comput. Mater. Sci.* **155**, 331–339 (2018).
18. Kim, K. et al. Deep-learning-based inverse design model for intelligent discovery of organic molecules. *npj Comput. Mater.* **4**, 67 (2018).
19. Raccuglia, P. et al. Machine-learning-assisted materials discovery using failed experiments. *Nature* **533**, 73 (2016).
20. Andrea, R., Sangid, M. D., Henry, P. & Wolfgang, L. Using machine learning and a data-driven approach to identify the small fatigue crack driving force in polycrystalline materials. *npj Comput. Mater.* **4**, 35 (2018).
21. Bhadeshia, H. K. D. H. Neural networks in materials science. *ISIJ Int.* **39**, 966–979 (1999).
22. Datta, S. & Banerjee, M. K. Mapping the input–output relationship in HSLA steels through expert neural network. *Mater. Sci. Eng. A* **420**, 254–264 (2006).
23. Reddy, N. S., Krishnaiah, J., Young, H. B. & Lee, J. S. Design of medium carbon steels by computational intelligence techniques. *Comput. Mater. Sci.* **101**, 120–126 (2015).
24. Ozerdem, M. S. & Kolukisa, S. Artificial neural network approach to predict the mechanical properties of Cu-Sn-Pb-Zn-Ni cast alloys. *Mater. Des.* **30**, 764–769 (2009).
25. Lu, W., Xiao, R., Yang, J., Li, H. & Zhang, W. Data mining-aided materials discovery and optimization. *J. Mater.* **3**, 191–201 (2017).
26. Wu, Y. et al. Correlations between Microstructures and properties of Cu-Ni-Si-Cr alloy. *Mater. Sci. Eng. A* **731**, 403–412 (2018).
27. Wang, W. et al. Correlation between microstructures and mechanical properties of cryorolled CuNiSi alloys with Cr and Zr alloying. *Mater. Charact.* **144**, 532–546 (2018).
28. Zhao, Z., Xiao, Z., Li, Z., Ma, M. & Dai, J. Effect of magnesium on microstructure and properties of Cu-Cr alloy. *J. Alloy. Compd.* **752**, 191–197 (2018).
29. Guo, F. et al. Study of rare earth elements on the physical and mechanical properties of a Cu-Fe-P-Cr alloy. *Mater. Sci. Eng. B* **147**, 1–6 (2008).
30. Watanabe, C., Takeshita, S. & Monzen, R. Effects of small addition of Ti on strength and microstructure of a Cu-Ni-Si alloy. *Metall. Mater. Trans. A* **46**, 2469–2475 (2015).
31. Monzen, R. & Watanabe, C. Microstructure and mechanical properties of Cu-Ni-Si alloys. *Mater. Sci. Eng. A* **483**, 117–119 (2008).
32. Watanabe, H., Kunimine, T., Watanabe, C., Monzen, R. & Todaka, Y. Tensile deformation characteristics of a Cu-Ni-Si alloy containing trace elements processed by high-pressure torsion with subsequent aging. *Mater. Sci. Eng. A* **730**, 10–15 (2018).
33. Keogh, E. & Mueen, A. *Encyclopedia of Machine Learning*. Ch. 129 (Springer US, Boston, MA, 2017).
34. Liu, F. et al. Microstructure and properties of low concentration Cu-Ni-Si alloy under different solution, aging temperature and cold deformation. *Chin. J. Rare Met.* **42**, 356–361 (2018).
35. Wang, W. et al. Effects of Cr and Zr additions on microstructure and properties of Cu-Ni-Si alloys. *Mater. Sci. Eng. A* **673**, 378–390 (2016).
36. Caron, R. N. & Breedis, J. F. Multipurpose copper alloys with moderate conductivity and high strength. U.S. Patent No. 4594221 (1986).
37. Suzuki, S. et al. Improvement in strength and electrical conductivity of Cu-Ni-Si alloys by aging and cold rolling. *J. Alloy. Compd.* **417**, 116–120 (2006).
38. Wang, H. et al. Improvement in strength and thermal conductivity of powder metallurgy produced Cu-N-Si-Cr alloy by adjusting Ni/Si weight ratio and hot forging. *J. Alloy. Compd.* **633**, 59–64 (2015).
39. Kim, Y. et al. Effect of heat treatment on precipitation behavior in a Cu-Ni-Si-P alloy. *J. Mater. Sci.* **21**, 1357–1362 (1986).
40. Zhang, Y. et al. Microstructure and precipitate's characterization of the Cu-Ni-Si-P alloy. *J. Mater. Eng. Perform.* **25**, 1336–1341 (2016).
41. Morozova, A., Mishnev, R., Belyakov, A. & Kaibyshev, R. Microstructure and properties of fine grained Cu-Cr-Zr alloys after thermo-mechanical treatments. *Rev. Adv. Mater. Sci.* **54**, 56–92 (2018).
42. Takagawa, Y., Tsujiuchi, Y., Watanabe, C., Monzen, R. & Tsuji, N. Improvement in mechanical properties of a Cu-2.0 mass% Ni-0.5 mass% Si-0.1 mass% Zr alloy by combining both accumulative roll-bonding and cryo-rolling with aging. *Mater. Trans.* **54**, 1–8 (2013).
43. Wang, Y. et al. Effects of deep cryogenic treatment on the solid-state phase transformation of Cu-Al alloy in cooling process. *Phase Transit.* **85**, 650–657 (2012).
44. He, S., Jiang, Y., Xie, J., Li, Y. & Yue, L. Effects of Ni content on the cast and solid-solution microstructures of Cu-0.4 wt% Be alloys. *Int. J. Miner. Metall. Mater.* **25**, 641–651 (2018).
45. Palavar, O., Özyürek, D. & Kalyon, A. Artificial neural network prediction of aging effects on the wear behavior of IN706 superalloy. *Mater. Des.* **82**, 164–172 (2015).
46. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).
47. Goh, A. T. C. Back-propagation neural networks for modeling complex systems. *Artif. Intel. Eng.* **9**, 143–151 (1995).
48. Xia, X. et al. An artificial neural network for predicting corrosion rate and hardness of magnesium alloys. *Mater. Des.* **90**, 1034–1043 (2016).
49. Sha, W. & Edwards, K. L. The use of artificial neural networks in materials science based research. *Mater. Des.* **28**, 1747–1752 (2007).