

ARTICLE OPEN



Divide and conquer: Machine learning accelerated design of lead-free solder alloys with high strength and high ductility

Qinghua Wei^{1,4}, Bin Cao ^{1,2,4}, Hao Yuan^{1,4}, Youyang Chen¹, Kangdong You¹, Shuting Yu¹, Tixin Yang¹, Ziqiang Dong¹ and Tong-Yi Zhang ^{1,2,3}

The attainment of both high strength and high ductility is always the goal for structure materials, because the two properties generally are mutually competing, called strength-ductility trade-off. Nowadays, the data-driven paradigm combined with expert domain knowledge provides the state-of-the-art methodology to design and discovery for structure materials with high strength and high ductility. To enhance both strength and ductility, a joint feature is proposed here to be the product of strength multiplying ductility. The strategy of “divide and conquer” is developed to solve the contradictory problem, that material experimental data of mechanical behaviors are, in general, small in size and big in noise, while the design space is huge, by a newly developed data preprocessing algorithm, named the Tree-Classifier for Gaussian Process Regression (TCGPR). The TCGPR effectively divides an original dataset in a huge design space into three appropriate sub-domains and then three Machine Learning (ML) models conquer the three sub-domains, achieving significantly improved prediction accuracy and generality. After that the Bayesian sampling is applied to design next experiments by balancing exploitation and exploration. Finally, the experiment results confirm the ML predictions, exhibiting novel lead-free solder alloys with high strength high ductility. Various material characterizations were also conducted to explore the mechanism of high strength and high ductility of the alloys.

npj Computational Materials (2023)9:201; <https://doi.org/10.1038/s41524-023-01150-0>

INTRODUCTION

Along with the fast development of artificial intelligence and machine learning (ML), the data-driven paradigm of materials informatics, which unifies the knowledge learned from experiments, theory, computations, and simulations, is rapidly becoming popular in materials science and engineering^{1–8}. The data-driven paradigm combined with expert domain knowledge provides state-of-the-art methodologies to understand and predict complex behaviors of materials and great achievements have been obtained in materials informatics^{9–16}. Materials data, especially experimental data of mechanical behaviors of materials, are small in size, whereas the influence factors, including testing conditions (or service environments), material compositions and microstructures, specimen size, etc., are considerably large, meaning that the dimensions of feature space, called search space also, are extremely high. Besides, some experimental data might scatter greatly. Wang et al.¹⁷ illustrate that domain knowledge guided statistic learning is powerful in tackle small data, where the knowledge of fracture mechanics and material strength theory suggest four hypotheses to guide the development of the size-dependent normal distribution model and the size- and pre-notch length-dependent normal distribution model. As a result, they successfully estimate the physical properties of fracture toughness and fracture process zone size. Cao et al.¹⁸ proposed a domain knowledge-guided interpretive machine learning strategy and demonstrated it by studying the oxidation behavior of ferritic-martensitic steels in supercritical water. A ML algorithm of Tree-Classifier for Linear Regression (TCLR) is developed which effectively captures the linear correlation between compositions, testing environments and oxidation behaviors from sparse data with high dimensions. Consequently, a generalized Arrhenius

oxidation formula is accomplished with very high prediction accuracy and wide generality. Wei et al.¹⁹ adopted the same strategy¹⁸ to discover high interpretive formula describing the high temperature oxidation behavior of FeCrAlCoNi-based high entropy alloys (HEAs). The TCLR algorithm was used to extract the spectrums of activation energy Q and time exponent m from the complex and high dimensional feature space, which automatically gives the spectrum of pre-factor. The three spectrums are assembled by using the element features, which leads to a general and interpretive formula with high prediction accuracy of the determination coefficient $R^2 = 0.971$. The role of each chemical element in the high temperature oxidation behavior is analytically illustrated in the three spectrums, thereby the discovered interpretative formula provides a guidance to the inverse design of HEAs against high temperature oxidation. The present work follows the similar methodology to develop a new data preprocessing algorithm of Tree-Classifier for Gaussian Process Regression (TCGPR) to tackle sparse nonlinear data. Both TCGPR and TCLR take the advantage of Tree-Classifier. TCLR uses a linear relationship between response and feature as the data slitting criterion so that each of leaves contains only good linear relationship data, as demonstrated in the previous works^{18,19}. In addition, TCGPR also takes that advantage of Gaussian Process Regression with Gaussian radial kernel function, where the length scale plays an essential role in the identification of outliers. The Global Gaussian Messy Factor (GGMF) has been proposed in TCGPR as the data partition criterion so that TCGPR can partition data into different distributions and identify outliers.

For the application of a material, the material must possess optimal balanced multiple properties. For structural materials, strength and ductility are two basic mechanical properties and

¹Materials Genome Institute, Shanghai University, Shanghai 200444, China. ²Advanced Materials Thrust, Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511400 Guangdong, China. ³Guangzhou Municipal Key Laboratory of Materials Informatics, Guangzhou 511400 Guangdong, China. ⁴These authors contributed equally: Qinghua Wei, Bin Cao, Hao Yuan. email: zqdong@shu.edu.cn; zhangty@shu.edu.cn

both are usually repellent each other, meaning that increasing strength is commonly accompanied with a decrease in ductility, and this behavior is called strength-ductility trade-off. Optimizing two or more targeted properties simultaneously is demanded in developing new materials and called multi-task optimization in ML. Wei et al.²⁰ used three multi-objective ML algorithms of MultiTaskLasso, Random forest, and Artificial neural network to predict simultaneously four mechanical properties of steels, including fatigue strength, tensile strength, fracture strength, and hardness. Xie's group (Wang et al.²¹, 2019, Jiang et al.²², Zhang et al.²³) developed a machine learning design system with two ML modules, (1) composition (prediction) design (P2C) module with properties as inputs and (2) property prediction (C2P) module with composition as inputs. With desired multiple properties, candidate alloy composition is generated through the P2C module and then the P2C generated composition is used as inputs into the C2P module to predict the properties. Employing the machine learning design system, Jiang et al.²² successfully designed and fabricated high-strength and high-toughness aluminum alloys and Zhang et al.²³ investigated copper alloys with high strength and high electrical conductivity. Using Pareto plot in the strength and ductility coordinate system, Chen et al.⁶ proposed an active learning strategy to perform multi-task optimization of strength and ductility of Mg alloys. Their design strategy is to select the best trade-off solution by two approaches in the Pareto plot of strength and ductility. The approach I minimizes the angle between the target vector of strength and ductility and an optimal predicted point on Pareto front and the approach II minimizes the distance between target point and an optimal predicted point on the Pareto front, which is a set of solutions giving the potential best trade-off between the competing properties and forms a characteristic boundary in the Pareto plot. Multi-objective optimization technique²⁴ was proposed to design the Pareto front of organic dielectric polymers. Following the Pareto front approach, we propose here the product of strength multiplying ductility as a joint feature such that optimizing the joint feature optimizes the Pareto front. We apply the developed strategy, including joint feature, TCGPR, Bayesian sampling, and experiment, to study lead-free solders.

Historically, Pb-Sn solders have been widely used in the electronics industry due to low melting point, excellent solderability, good electrical conductivity, and low cost^{25–29}. However, Lead is a toxic metal, pollutes the environment and endangers human health³⁰ and thus, many countries have banned Pb-Sn solders^{31–33}. Currently, Sn-Ag-Cu(SAC) based alloys, near eutectic composition, are widely used as lead-free solders due to low melting points and excellent solderabilities^{34–37}. The microelectronic and electronic industries demand even better performance than currently used SAC solder alloys. How to further enhance the performance lead-free solders is a great challenge and the present work endeavors to confront this challenge by using the domain knowledge guided active learning ML.

Solder alloys are generally used for joint connections. Hence, besides the electronic conductivity, the mechanical properties of the solder alloys should be primarily considered, which include their strength, ductility, toughness, creep, fatigue resistance, etc. Overall, the intrinsic mechanical properties of solder alloys should be comparable or better than those of Pb-Sn solder alloys to meet the application requirements in industries, within which strength and ductility are the two essential mechanical properties. As described above, how to enhance both strength and ductility simultaneously has to solve the trade-off problem. Considerable efforts have been devoted to optimizing the mechanical properties of SAC alloys. Besides thermomechanical treatments, alloying is an effective method to improve the mechanical properties of SAC alloys. There are two basic strengthening mechanisms that alloying elements can play. The first strengthening mechanism is the solid solution hardening, when alloying atoms are dissolved in the matrix phase.

The second strengthening mechanism is the precipitation hardening, when alloying atoms interact with matrix solvent atoms and form intermetallic compound (IMC). Usually, small and uniformly distributed IMC particles have the best strengthening performance. That is why various alloying elements have been added to the basic SAC alloy (Sn-3.8Ag-0.7Cu) to explore potential improvement of mechanical properties. The Ag₃Sn and Cu₃Sn IMCs can be formed in SAC alloys to strengthen the alloys. In addition, the pre-existence of Cu in SAC alloy could also inhibit the dissolution of Cu from a Cu-based substrate into the solder joint during welding. There are many reports in the literature on the effects of alloying elements on the solder performance. El-Daly et al.³⁸ found that adding Bi can significantly improve the tensile strength and enhance the creep resistance of brazing materials owing to solid solution strengthening and/or precipitation strengthening. Alloying suitable amount of element In could lead to a more uniform distribution of IMC precipitations in the matrix of brazing material and hence improve the tensile strength³⁹. The addition of Zn might refine the Ag₃Sn and Cu₆Sn₅ IMCs and form the (Cu, Ag)₅Zn₈ IMC, which strengthens the brazing material⁴⁰. Alloying Ti element can effectively refine the grain size of Sn-3.5Ag-0.5Cu brazing alloy and generate Ti₂Sn₃ IMC and hence strengthens the alloy by the grain refinement and precipitation⁴¹. Alloying Sb into Sn-3Ag-0.5Cu solder alloy strengthens the alloy by solid solution hardening and the precipitation hardening of the formation of Ag₃(Sn, Sb) and Cu₆(Sn, Sb)₅ IMCs⁴². Alloying Al element refines the β-Sn phase particles in SAC105 solder, inhibits the formation of Ag₃Sn and Cu₆Sn₅ IMCs, and forms new Ag₃Al and Al₂Cu IMCs, thereby significantly increasing the strength and decreasing the ductility⁴³. Alloying Ni atoms generate a more dense and stable (Cu, Ni)₆Sn₅ IMC interface layer between solder and Cu-based substrate, which becomes a barrier layer against the growth of brittle Cu₃Sn IMC and hence improves the drop resistance⁴⁴. However, big size IMCs are usually very brittle, meaning strengthening by the precipitation mechanism might impair the ductility when IMCs are big in size. How to control the amounts of alloying elements and the size and distribution of IMCs is the bottleneck problem in the optimal design and fabrication of lead-free solders with high strength and high ductility.

The present work studies the alloying effects on the strength and ductility of lead-free solders by investigating Sn_{bal}Ag_{3.8}Cu_{0.7}Bi_{x₁}Zn_{x₂}Sb_{x₃}In_{x₄}Ti_{x₅}Ni_{x₆}Al_{x₇}. The ranges and variation steps of each element are described in the method, which form a huge search space of 17931 samples. Initial 47 samples are designed, fabricated and tested by tensile tests, providing the initial sparse data. To tackle the sparse data, the present work proposes the divide-and-conquer strategy to accelerate the alloy design process. Figure 1 shows the framework of the present work for the alloy design. The initial data consist of alloy composition, ultimate tensile strength and fracture elongation. The product of ultimate tensile strength and fracture elongation is taken as the joint feature. The TCGPR algorithm is developed to divide the training data into three clusters, which accordingly selects only three subspaces from the original search space and enhances the ML prediction accuracy. The Bayesian sampling method is used to recommend the potential candidates for experiment and the experimental results show the improved mechanical properties, better than these in the training dataset and conventional Sn solders.

RESULTS

Experimental data

47 data of SAC387 based alloys, Sn_{bal}Ag_{3.8}Cu_{0.7}Bi_{x₁}Zn_{x₂}Sb_{x₃}3In_{x₄}Ti_{x₅}Ni_{x₆}Al_{x₇}, were generated from in-house experiments, which compositions in wt.% include fixed Ag content of 3.8 wt.% and Cu concentration of 0.7wt.%, 7 adjustable alloying elements, and balanced by Sn. The experimental 47 data of

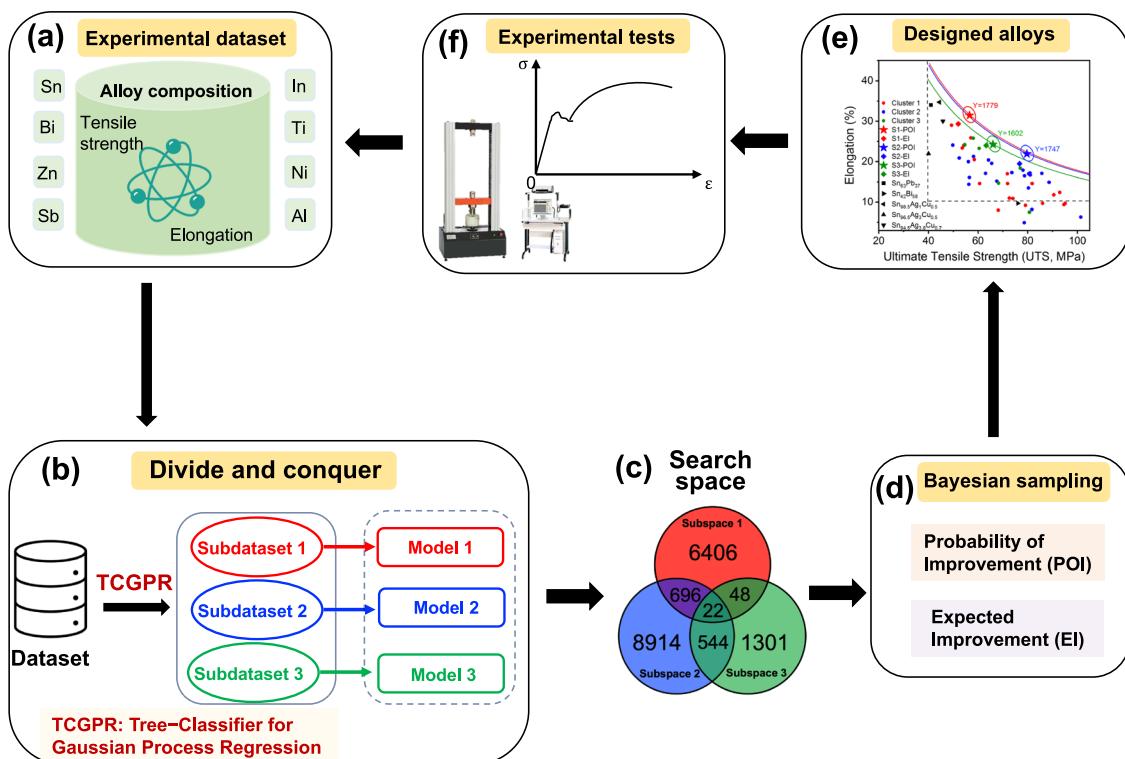


Fig. 1 Workflow. Designing framework for lead-free solder alloy with high strength and high ductility.

Table 1. The variation range and step size of alloying elements.

Composition	Sn	Ag	Cu	Bi	Zn	Sb	In	Ti	Ni	Al
Training data element range	Bal.	3.8	0.7	0–5	0–1	0–5	0–5	0–0.7	0–0.8	0–0.8
Total virtual sample element range	Bal.	3.8	0.7	0–5.5	0–1.2	0–5.5	0–5.5	0–0.8	0–0.9	0–0.9
Step size(wt. %)	--	--	--	0.5	0.2	0.5	0.5	0.1	0.1	0.1

SAC387 based alloys is given in Supplementary Information. Table 1 lists the variation range and step size in concentration of the 7 adjustable alloying elements in the initially tested samples and in the search space, the concentration of each alloying element in the initial dataset is extended outward by one step to form a search space of virtual 19232 samples. The ultimate tensile strength (UTS) is in units of MPa and fracture elongation is in percentage.

Data preprocessing and machine learning

To enhance the strength-ductility synergy, a joint objective is proposed here

$$Y = \sigma_u \times \epsilon_f \quad (1)$$

subject to $\sigma_u \geq \sigma_{u,0}$ and $\epsilon_f \geq \epsilon_{f,0}$

where σ_u and ϵ_f denotes the ultimate tensile strength (UTS) and the engineering strain at fracture, respectively, and $\sigma_{u,0}$ and $\epsilon_{f,0}$ are the minimal values preset by the designer. The present work uses $\sigma_{u,0} = 40$ MPa and $\epsilon_{f,0} = 10\%$. The product (joint objective) of strength and ductility provides an approximate estimate of, at least proportional to, the energy absorbed by a tested sample during the tensile test, which has a clear physical meaning. To maintain explicitly this physical meaning, the original physical

dimensions of σ_u and ϵ_f are used in the joint objective, and the joint objective is in unit of MPa.

The ML algorithms of Random Forest (RF), eXtreme Gradient Boosting (XGB), and GPR are used in the present work with the leave-one-out cross validation (LOOCV) and Pearson correlation coefficient R to evaluate the performance of ML models, and hereafter, LOOCV R represents the fitting goodness of a ML model. Figure 2 (a-d) show the predictive value versus the experimental value of the joint objective, where (a-c) indicate the predictions without using the TCGPR by the RF, XGB, and GPR, respectively, and (d) illustrates the predictions with using the TCGPR by three GPR models. Obviously, the divide and conquer TCGPR model significantly improves the LOOCV R . Specifically, the LOOCV R of RF, XGB and GPR on the original dataset are 0.25, 0.37 and 0.33 respectively, whereas the LOOCV R of three GPR models on Clusters 1, 2, and 3 are 0.91, 0.81 and 0.78, respectively, as shown in Fig. 2. Figure 3 (a-d) show the LOOCV prediction error distributions of the joint objective (Y) for the models with and without the use of TCGPR. The mean value (μ) of predicted errors of models with or without the use of TCGPR are all small and close to zero, ranging from 5.75 (MPa) to 18.93 (MPa). The utilization of TCGPR decreases considerably the standard deviation (σ) of predicted errors, from 286.59 (MPa) for RF, 286.52 (MPa) for XGB, and 264.53 (MPa) for GPR to 157.89 (MPa), as shown in Fig. 3. The

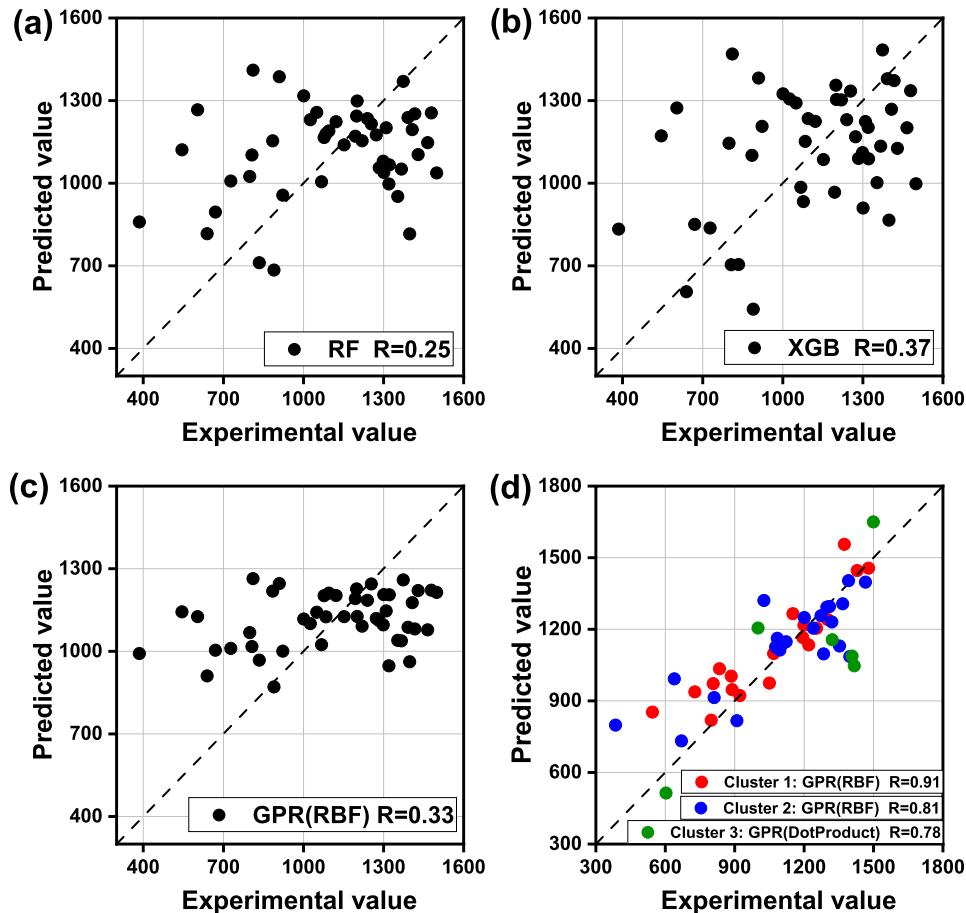


Fig. 2 The predictive value versus the experimental value of the joint objective (MPa). **a–c** without using TCGPR, **(a)** RF model, **(b)** XGB model, **(c)** GPR model, **(d)** with using TCGPR, three GPR models.

result indicates the TCGPR enhances the correlation among the data points within each cluster in comparison to that in the whole original dataset. The comparison results of TCGPR, K-means and hierarchical clustering algorithms on the dataset of 47 SAC387-based alloys is summarized in Supplementary Table 1. Both K-means and hierarchical clustering failed to improve the prediction accuracy after dividing the original dataset into clusters. During an epoch of TCGPR, one subdataset is selected from all possible subdatasets with the minimal global Gaussian messy factor (GGMF), corresponding to the maximal correlation among the data. The GGMF is calculated based on the GPR model with LOOCV and using the information of feature variables and the target variable. The minimum of GGMF naturally leads to a high LOOCV R of GPR prediction, which is the key reason of the TCGPR performing better than K-means and Hierarchical clustering algorithms.

The influence of the kernel functions on the model performance is also investigated by using the three kernels of Gaussian radial basis function (RBF) kernel, Exp-Sine-Squared kernel and Dot-Product kernel. The LOOCV R selects the optimal kernel functions are RBF, RBF, RBF, and Dot-Product, respectively, for the four datasets of original dataset, Cluster 1, Cluster 2, and Cluster 3. Supplementary Fig. 2a–c in the Supplementary Information show the LOOCV predictions of the three GPR models with the RBF, Exp-Sine-Squared and Dot-Product kernels versus the experimental values of the joint objective (Y) in three clusters, respectively.

Figures 2 and 3 indicate that the whole original data do not follow a same distribution and hence it is hard to use a GPR model to make the prediction. Due to TCGPR, the initial dataset is divided

into three clusters and each cluster corresponds a search subspace, where the tested samples are more correlated, as shown in Fig. 2, which also implies that there are less correlations among the three clusters. The physical meaning of more correlated tested samples in a cluster is that the correlation between the alloying elements and the joint objective under the invariant experimental conditions follows the same statistical distribution, which can be learnt by a used GPR. On the other hand, the tested samples in the original dataset are less correlated because the joint objective and alloying elements do not follow the same statistical distribution, which cannot be learnt by a used GPR. Considering the complicated correlations in the original dataset, divide and conquer is an excellent strategy and TCGPR is developed to implement the divide task and hence each cluster can be conquered by a GPR model.

The cluster labels of 47 experimental data are given in Supplementary Table 5, where there are 19, 22, and 6 data in Clusters 1, 2, and 3, respectively. Three search subspaces are constructed based on the composition in the three clusters. As described above for original dataset, the search feature space is defined by outward expanding one step of each element concentration. The steps for the alloying elements Bi, Zn, Sb, In, Ti, Ni and Al are 0.5 wt.%, 0.2 wt.%, 0.5 wt.%, 0.5 wt.%, 0.1 wt.%, 0.1 wt.% and 0.1 wt.%, respectively. The same expansion approach is conducted to set the search space of each cluster, which causes the overlap if the experimental data in different clusters are close. Table 2 lists the variation range and step size of alloying elements in the three search subspaces, indicating that there is no Al element in subspace 3, because no Al element in the cluster 3 of

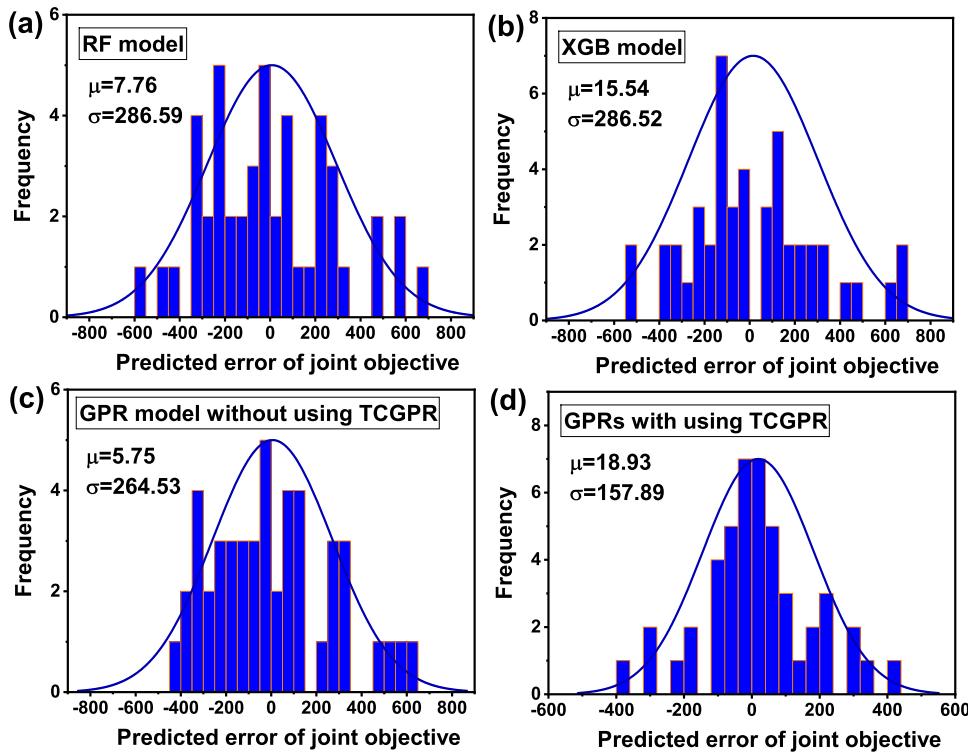


Fig. 3 The LOOCV prediction error histogram of the joint objective (Y). a–c without using TCGPR, (a) RF model, (b) XGB model, (c) GPR model, and (d) the three GPR models on the three clusters grouped by TCGPR, where the prediction error, mean and standard deviation are all in the units of MPa.

Table 2. The variation range and step size of alloying elements.

Composition	Sn	Ag	Cu	Bi	Zn	Sb	In	Ti	Ni	Al
Subspace 1	Bal.	3.8	0.7	0–5.5	0–0.6	0–5.5	0–5.5	0–0.8	0–0.9	0–0.9
Subspace 2	Bal.	3.8	0.7	0–5.5	0–1.2	0–5.5	0–5.5	0–0.8	0–0.9	0–0.9
Subspace 3	Bal.	3.8	0.7	0–3	0–0.8	0–3	0–1.5	0–0.2	0–0.5	--
Step size(wt. %)	--	--	--	0.5	0.2	0.5	0.5	0.1	0.1	0.1

the tested samples. Figure 4 indicates the numbers of virtual samples in the three subspaces, which are partially jointed among themselves. The TCGPR reduces the total number of virtual samples from 19232 to 17931, as shown in Fig. 4, the data correlation is greatly enhanced in each subspace.

Bayesian sampling results

Gaussian process regression provides predictions in terms of the mean and variance of all virtual samples in the search subspaces. Then, with the joint objective $Y = \sigma_u \times \varepsilon_f$ subject to $\sigma_u \geq 40 and $\varepsilon_f \geq 10\%$, the acquisition function of probability of improvement (POI)^{45–47} and expected improvement (EI)⁴⁸ are used to combine the terms of the predicted mean and variance in three search subspaces. Finally, the Bayesian sampling recommended six samples from search subspaces for next experiments by balancing exploration and exploitation, the recommended samples are listed in Table 3.$

Experimental results of the designed alloy samples

The experimental results of six designed alloys are listed in Table 3. Figure 5a shows the comparison of the predictive value and the experimental value for the comprehensive properties (Y), indicating that the experimental results confirm basically the

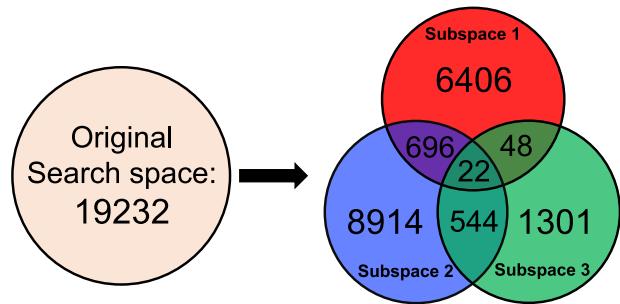


Fig. 4 Search space. The TCGPR divides the initial search space into three subspaces, where red, blue and green represent subspaces 1, 2, and 3, respectively.

predictions correspondingly. For each recommended alloy, one representative stress-strain curve was plotted and the stress-strain curve of the highest comprehensive property (Y) in the training data cluster was also plotted for comparison. Figure 5b–d show these stress-strain curves, where the solid stress-strain curves denote for the representative recommended alloys and the dash stress-strain curves are for the highest comprehensive

Table 3. Six alloys recommended by Bayesian sampling and experimental results.

Alloy No.	Composition	UTS (MPa)	Elongation (%)	Y(MPa)	Improvement (%)
S1-POI	$\text{Sn}_{94.8}\text{Ag}_{3.8}\text{Cu}_{0.7}\text{Bi}_{0.5}\text{Zn}_{0.2}$	56.57 ± 4.14	31.45 ± 0.99	1779 ± 142	20.31
S1-EI	$\text{Sn}_{94.7}\text{Ag}_{3.8}\text{Cu}_{0.7}\text{Bi}_{0.5}\text{Zn}_{0.2}\text{Ti}_{0.1}$	52.00 ± 0.96	29.34 ± 2.59	1526 ± 123	3.17
S2-POI	$\text{Sn}_{90.8}\text{Ag}_{3.8}\text{Cu}_{0.7}\text{Bi}_3\text{In}_{1.5}\text{Ti}_{0.2}$	79.68 ± 1.23	21.93 ± 0.50	1747 ± 166	19.28
S2-EI	$\text{Sn}_{90.5}\text{Ag}_{3.8}\text{Cu}_{0.7}\text{Bi}_{2.5}\text{In}_{1.5}\text{Ti}_{0.3}\text{Zn}_{0.2}\text{Sb}_{0.5}$	76.68 ± 2.00	19.50 ± 1.83	1495 ± 112	2.07
S3-POI	$\text{Sn}_{92.6}\text{Ag}_{3.8}\text{Cu}_{0.7}\text{Zn}_{0.6}\text{Sb}_{2.5}\text{In}_{0.5}\text{Ti}_{0.2}\text{Ni}_{0.1}$	66.04 ± 1.47	24.26 ± 3.33	1602 ± 180	6.92
S3-EI	$\text{Sn}_{92.3}\text{Ag}_{3.8}\text{Cu}_{0.7}\text{Zn}_{0.6}\text{Sb}_{2}\text{In}_{0.5}\text{Ni}_{0.1}$	63.24 ± 3.98	24.00 ± 3.00	1518 ± 158	1.28

Notes: In the Alloy No., the first part of S1, S2 and S3 indicate the search subspaces 1, 2, and 3, respectively; while the second part of POI and EI denote the acquisition functions POI and EI, respectively.

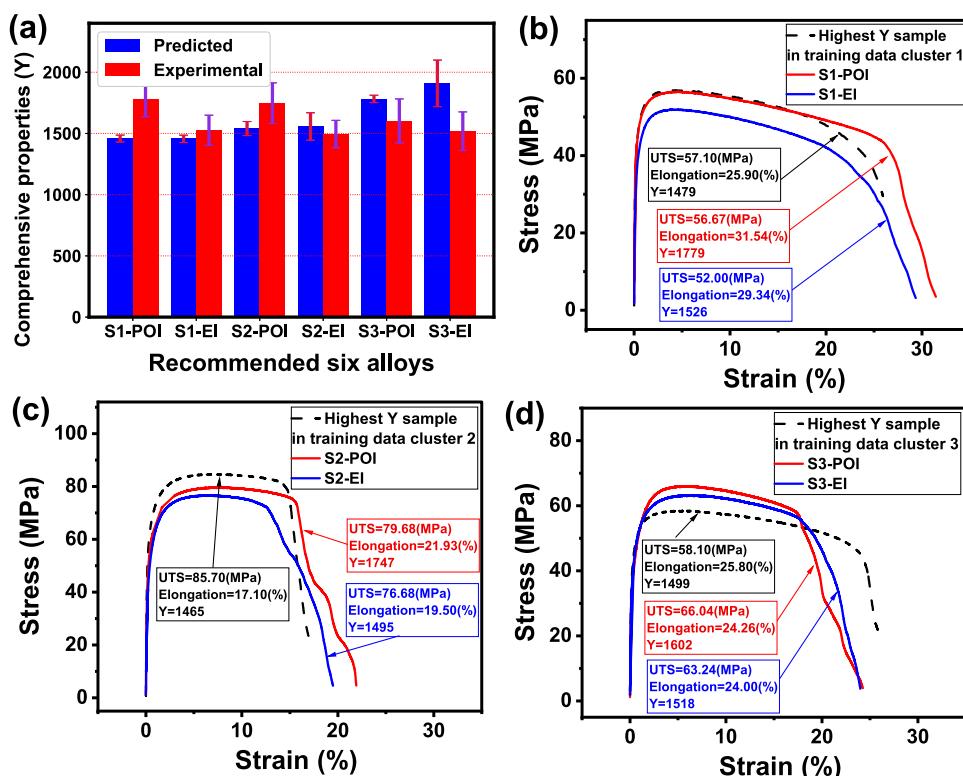


Fig. 5 The comparison of the predictive value and the experimental value for the comprehensive properties (Y). **a** The experimental measured and ML predicted values of comprehensive property (MPa) for six recommended alloys, where the error bar denotes the variance; Stress-strain curves for **(b)** cluster 1, **(c)** cluster 2 and **(d)** cluster 3. The solid stress-strain curves denote for the representative recommended alloys and the dash stress-strain curves are for the highest comprehensive property in the training data clusters.

properties (Y) in the training data clusters. The highest comprehensive properties in the training cluster 1, cluster 2 and cluster 3 are 1479, 1465 and 1499, respectively. It is obvious that the experimental results of comprehensive mechanical properties on the six recommended samples are all higher than the best samples in the corresponding training clusters. Compared with the best sample in the first training data cluster, the comprehensive mechanical properties of sample S1-POI and sample S1-EI are improved by 20.31% and 3.17%, respectively. Similarly, the comprehensive mechanical properties of S2-POI, sample S2-EI, sample S3-POI and sample S3-EI are also improved by 19.28%, 2.07%, 6.92% and 1.28% respectively, as listed in Table 3. The results indicate that the comprehensive mechanical properties (Y) of the three alloys recommended by acquisition function POI (sample S1-POI, S2-POI and S3-POI) are all greater than those recommended by acquisition function EI (sample S1-

EI, S2-EI and S3-EI). In the present study, the POI acquisition strategy performs better than the EI acquisition strategy.

Figure 6 shows all experimental data and the Pareto front, where the sample S1-POI, sample S2-POI and sample S3-POI are on the Pareto front. Obviously, the three samples exhibit the better mechanical properties in terms of strength and ductility than all samples in the training datasets, implying the effectiveness of the approach in optimizing the strength-ductility synergy. Especially, sample S1-POI is the alloy with a good ductility, sample S2-POI represents the solder alloy of the high strength with a compromised ductility and sample S3-POI is the alloy with a balanced strength and ductility. Subsequently, the experimental characterization and mechanism exploration are performed on the three samples S1-POI, S2-POI and S3-POI.

Experimental characterization and mechanism exploration

The microstructures and fracture morphologies of the three novel alloys (S1-POI, S2-POI, and S3-POI) were examined and the strengthening mechanisms were discussed below:

(1) High ductility solder alloy: $\text{Sn}_{94.8}\text{Ag}_{3.8}\text{Cu}_{0.7}\text{Bi}_{0.5}\text{Zn}_{0.2}$ (S1-POI): the microstructure of alloy S1-POI is shown in Fig. 7a, b, and its element distribution map and the results of elemental quantitative analysis are shown in Supplementary Fig. 3 and Supplementary Table 2, respectively. The alloy consists of two phases of the Sn-based solid solution phase and the eutectic phase. Sn distributes evenly across the alloy, while Ag and Cu are mainly enriched in the Intermetallic Compounds (IMC) (Ag_3Sn , Cu_6Sn_5) precipitated at

the eutectic regions. The sample shows a high ductility which could be attributed to the relative lower contents of the alloying elements added. The solubility of Bi and Zn in Sn are about 1.0 ~ 1.8 wt. %^{49–53} and 0.5 ~ 0.6 wt. %⁵⁴, respectively. At current concentration level, both Bi and Zn could be completely dissolved into Sn matrix, which could improve the strengthen of the solder alloy due to the solid solution strengthening effect. Meanwhile, adding Bi and Zn could refine the microstructure^{40,55}, which is beneficial to improving the ductility of solder alloy. The fracture morphologies of alloy S1-POI after the tensile tests are shown in Fig. 8a, b, indicating the typical ductile failure mode. As shown in Fig. 8b, dimples and voids appear across the fractured surface, demonstrating the high ductility. The fracture morphologies are consistent with the tensile test results of the largest elongation of $31.45 \pm 0.99\%$.

(2) High strength solder alloy: $\text{Sn}_{90.8}\text{Ag}_{3.8}\text{Cu}_{0.7}\text{Bi}_{3}\text{In}_{1.5}\text{Ti}_{0.2}$ (S2-POI): alloy S2-POI demonstrates an excellent comprehensive mechanical property with a tensile strength of about 80 MPa and an elongation of 22%. The enhanced strength could be attributed to the synergistic strengthening effects caused by solution, precipitation, and grain refinement. The microstructure of alloy S2-POI is shown in Fig. 7c, d. The elemental distribution map and corresponding quantitative analysis results are shown in Supplementary Fig. 4 and Supplementary Table 3, respectively. Generally, adding Bi into Sn alloy could effectively enhance the strength of solder alloy due to solid solution strengthening effect^{50,51}. As mentioned above, the solubility of Bi in Sn is about 1.0–1.8 wt%. When the content of Bi reaches 3 wt. %, excess Bi atoms precipitate out from Sn matrix, as shown in Fig. 7d. The element In can also be dissolved into Sn matrix and cause solution strengthening effect and could replace the Sn atoms in Ag_3Sn and Cu_6Sn_5 , forming $\text{Ag}_3(\text{Sn}, \text{In})$ and $\text{Cu}_6(\text{Sn}, \text{In})_5$. As shown in Fig. 7d, Ti could react with Sn and In to form the strip-shape intermetallic phase $\text{Ti}_2(\text{Sn}, \text{In})_3$, which could further strengthen the solder alloy⁴¹. The presence of the various IMCs and Bi precipitates could effectively block the dislocation movements and hence improve the tensile strength. However, the large IMCs and Bi precipitates might impair the ductility of the solder alloy since cracks could be initiated at the interface between the large precipitates and Sn

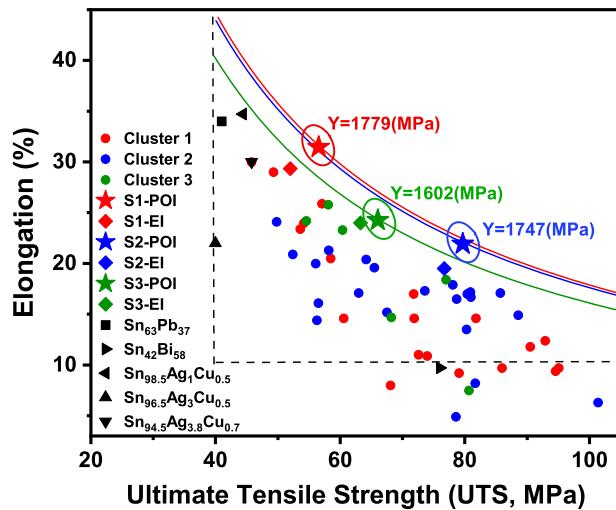


Fig. 6 Experimental validation results. The three pentagons and three rhombuses are the alloys designed by machine learning, and three curves are the Y contours, and five black icons denote the commercially available lead-free solder alloys for comparison.

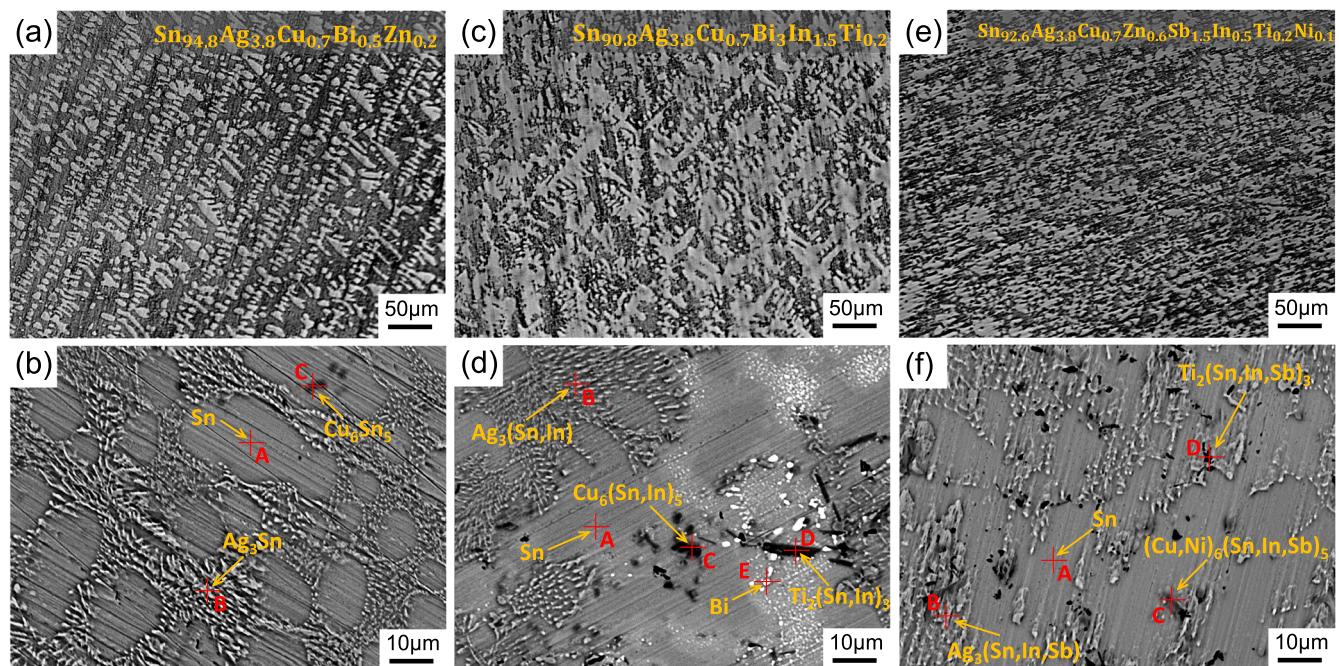


Fig. 7 Microstructures of the three novel alloys (S1-POI, S2-POI, and S3-POI). Optical microstructure images of alloys: (a) S1-POI, (c) S2-POI and (e) S3-POI; Backscattered electron (BSE) images of alloys: (b) S1-POI, (d) S2-POI and (f) S3-POI.

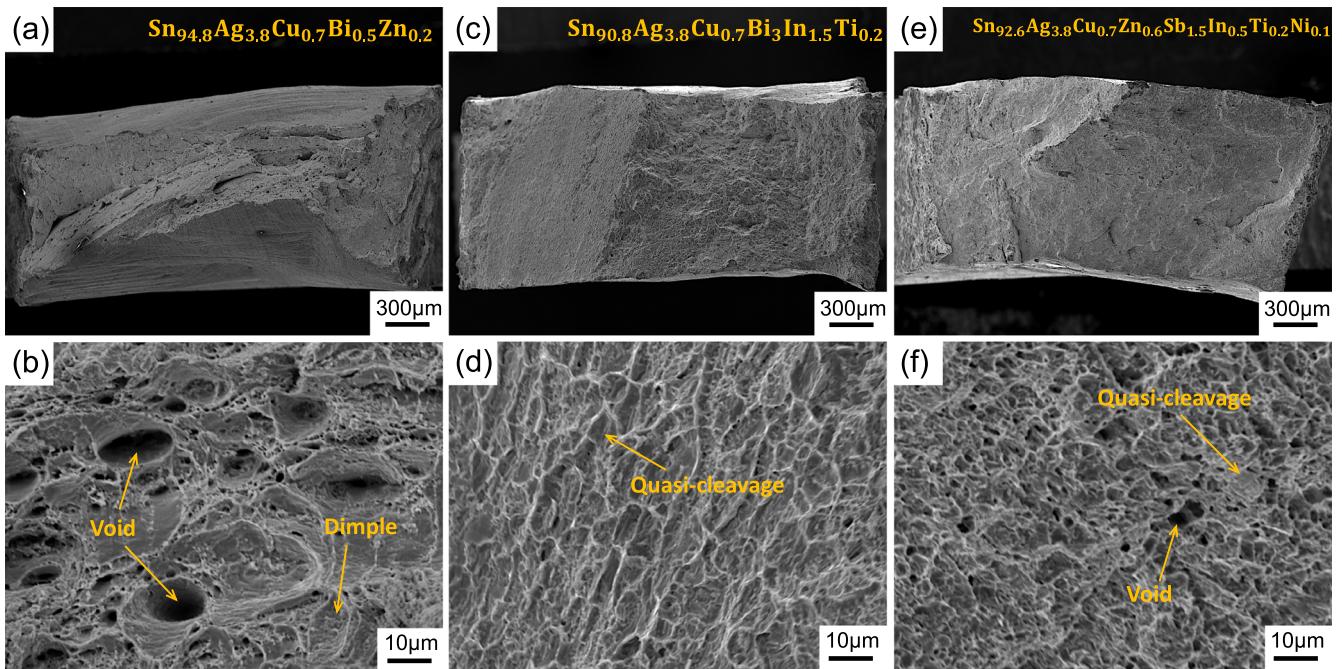


Fig. 8 Fracture morphologies. SEM images (Secondary electron image) of the tensile tested alloys: **a, b** S1-POI, **c, d** S2-POI and **e, f** S3-POI.

matrix. The fracture morphologies of the tested alloy S2-POI are shown in Fig. 8c, d. Alloy S2-POI shows a flatted fracture surface with the presence of quasi-cleavages, which indicates a lower ductility compared to alloy S1-POI.

(3) Solder alloy with the balanced mechanical properties: Sn_{92.6}Ag_{3.8}Cu_{0.7}Zn_{0.6}Sb_{1.5}In_{0.5}Ti_{0.2}Ni_{0.1} (S3-POI): this alloy is Bi-free and possesses more balanced mechanical properties regarding the strength and ductility. Other alloying elements show less strengthening effect than that of Bi. The microstructure of S3-POI is shown in Fig. 7e, f. The elemental distribution map is shown in Supplementary Fig. 5 and the elemental quantitative analysis results are listed in Supplementary Table 4. As shown in Fig. 7e, the microstructure of Sample S3-POI is much finer than those in S1-POI and S2-POI. The refined microstructure should be one key reason for the more balanced properties. It was reported that adding Zn could promote the heterogeneous nucleation of β -Sn⁴⁰, which can refine the structure. Meanwhile, adding In and Sb could cause solid solution strengthening effect which improves the strength of the solder alloy. As shown in Fig. 7f, In and Sb can also replace the Sn atoms in Ag₃Sn, Cu₆Sn₅ and Ti₂Sn₃ to form Ag₃(Sn, In, Sb), Cu₆(Sn, In, Sb)₅ and Ti₂(Sn, In, Sb)₃. Therefore, the strength of solder alloy is further improved due to the precipitation hardening effect. The fracture morphologies of the alloy S3-POI are shown in Fig. 8e, f, indicating that the mixture of dimples and quasi-cleavages, which is the typical fracture morphologies of ductile/brittle mixed fracture mode, corresponding to the balanced mechanical properties. Overall, alloy S3-POI shows a finer microstructure than those of S1-POI and S2-POI, due to the alloying effects of the multiple elements, thereby leading to the well-balanced mechanical properties.

DISCUSSION

In conclusion, this study proposes the divide and conquer strategy of active learning for alloy design, including data collection, data preprocessing, ML model construction, composition design, and experimental verification, to promote the rational design of lead-free solder alloys and improve both strength and ductility simultaneously. The previous works^{18,19} illustrate that the TCLR algorithm can split data into leaves and in each leaf there is a little linear cluster with the

linear relationship between response and feature, which imply that the original dataset is composed by many little linear clusters. If an original dataset is consisted of few nonlinear clusters, the developed TCGPR algorithm is able to effectively divide the dataset into few nonlinear clusters. In the present work, TCGPR divides the original dataset into three clusters and three machine learning models are accordingly constructed to predict the alloy comprehensive mechanical property (Y), which is defined as the product of tensile strength and fracture elongation, and approximately represents the energy absorbed by the tested sample under uniaxial tensile tests. The alloy with high comprehensive mechanical property is designed by Bayesian sampling and the following experimental tests verify the predictions from the ML models. As a result, three lead-free solder alloys with high comprehensive mechanical properties are successfully discovered. The present study fruitfully demonstrates that the divide and conquer strategy, although each ML model is developed on a fairly small data cluster. This is because the active adaptive learning is able to work on small initial dataset. Furthermore, once new experimental results are added into the data clusters, each of the updated data clusters might expand its search front towards the improvement direction and the pre-set values of $\sigma_{u,0}$ and $\varepsilon_{f,0}$ will be enhanced accordingly, which will be done in future study. The divide and conquer strategy of active learning paves the avenue for the design and discovery of novel advanced materials.

METHODS

TCGPR algorithm

Tree-Classifier for Gaussian process regression (TCGPR, <https://github.com/Bin-Cao/TCGPR>) is a data preprocessing algorithm developed for identifying outliers and/or cohesive data, which has been proposed for the first time in this work. TCGPR identifies outliers via Sequential Forward Identification (SFI). The SFI starts from few cohesive data, identifies outliers, which maximizes the expected decrease (ED) of the global Gaussian massy factor (GGMF) with a preset criterion of fitting-goodness, by adding a batch of $p \geq 1$ data in each sequential through the raw dataset, called an epoch. After an epoch, raw data is divided into one cohesive subset and a rest subset. In the following epoch, the rest

subset processed by TCGPR is divided into cohesive and rest subsets again. The preprocessing is going on until the raw dataset is divided into a series of highly cohesive subsets and a final rest subset containing outliers only. The detailed working principle of TCGPR model is given in Supplementary Information.

Bayesian sampling

When considering uncertainty in data, especially in predictions by a ML model, one must balance exploration and exploitation. From the probabilistic analysis, a prediction from a ML model is usually assumed to follow a normal distribution, where the variance represents the uncertainty. Gaussian process regression (GPR) used here searches optimal candidates with uncertainty. The acquisition function POI⁴⁷ and EI⁴⁸ are used to recommend experiments from the GPR prediction with uncertainty by balancing exploration and exploitation, which are defined respectively by

$$POI = \phi\left(\frac{\mu_i - y^*}{\sigma_i}\right) \quad (2)$$

$$EI = (\mu_i - y^*)\phi\left(\frac{\mu_i - y^*}{\sigma_i}\right) + \sigma_i\varphi\left(\frac{\mu_i - y^*}{\sigma_i}\right) \quad (3)$$

where μ_i and σ_i are the mean and variance of a normal distribution, respectively, $\varphi(\cdot)$ and $\phi(\cdot)$ are the distribution density and cumulative function, respectively, and y^* is the current best target value in the dataset.

Experiment

The raw materials are Sn, Ag, Cu, Bi, In, Ti, Sb, Zn, Al and Ni with a purity 99.99%. The melting process was conducted using an induction furnace and the molten alloy was cast in a cuboid steel mold. With the cast, the tested samples were prepared with a dog bone shape with a size of 12 mm × 5 mm × 5 mm. Tensile tests at room temperature were conducted using a MTS universal material testing system with a constant strain rate of $3 \times 10^{-3} \text{ s}^{-1}$. Eight repeated tests were conducted for each recommend alloy and their mean and standard deviation were reported here.

DATA AVAILABILITY

All experimental data in the study are contained in [Supplementary Information](#).

CODE AVAILABILITY

The source code of TCGPR is available at: <https://github.com/Bin-Cao/TCGPR>.

Received: 15 February 2023; Accepted: 4 October 2023;

Published online: 23 October 2023

REFERENCES

- Xiong, J., Zhang, T. & Shi, S. Machine learning of mechanical properties of steels. *Sci. China Technol. Sci.* **63**, 1247–1255 (2020).
- Agrawal, A. & Choudhary, A. An online tool for predicting fatigue strength of steel alloys based on ensemble data mining. *Int. J. Fatigue* **113**, 389–400 (2018).
- Wang, J. H., Sun, S., He, Y. L. & Zhang, T. Y. Machine learning prediction of the hardness of tool and mold steels (in Chinese). *Sci. Sin. Technol.* **49**, 1148–1158 (2019).
- Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A. & Kim, C. Machine learning in materials informatics: recent applications and prospects. *npj Comput. Mater.* **3**, 54 (2017).
- Tong, Z., Wang, L., Zhu, G. & Zeng, X. Predicting twin nucleation in a polycrystalline Mg alloy using machine learning methods. *Metall. Mater. Trans. A* **50**, 5543–5560 (2019).
- Chen, Y. et al. Machine learning assisted multi-objective optimization for materials processing parameters: A case study in Mg. *alloy. J. Alloy. Compd.* **844**, 156159 (2020).
- Agrawal, A. & Choudhary, A. Perspective: materials informatics and big data: realization of the “fourth paradigm” of science in materials science. *APL Mater.* **4**, 053208 (2016).
- Xiong, J., Shi, S. Q. & Zhang, T. Y. A machine-learning approach to predicting and understanding the properties of amorphous metallic alloys. *Mater. Des.* **187**, 108378 (2020).
- Sanchez-Lengeling, B. & Aspuru-Guzik, A. Inverse molecular design using machine learning: generative models for matter engineering. *Science* **361**, 360–365 (2018).
- Raccuglia, P. et al. Machine-learning-assisted materials discovery using failed experiments. *Nature* **533**, 73–76 (2016).
- Rickman, J. M., Lookman, T. & Kalinin, S. V. Materials informatics: from the atomic-level to the continuum. *Acta Mater.* **168**, 473–510 (2019).
- Dong, Z. et al. Machine learning-assisted discovery of Cr, Al-containing high-entropy alloys for high oxidation resistance. *Corros. Sci.* **220**, 111222 (2023).
- Rickman, J. M. et al. Materials informatics for the screening of multi-principal elements and high-entropy alloys. *Nat. Commun.* **10**, 2618 (2019).
- Su, Y., Fu, H., Bai, Y., Jiang, X. & Xie, J. Progress in materials genome engineering in China. *Acta Metall. Sin.* **56**, 1313–1323 (2020).
- Zhang, H. et al. Dramatically enhanced combination of ultimate tensile strength and electric conductivity of alloys via machine learning screening. *Acta Mater.* **200**, 803–810 (2020).
- Wen, C. et al. Machine learning assisted design of high entropy alloys with desired property. *Acta Mater.* **170**, 109–117 (2019).
- Wang, J. H., Jia, J. N., Sun, S. & Zhang, T. Y. Statistical learning of small data with domain knowledge—sample size-and pre-notch length-dependent strength of concrete. *Eng. Fract. Mech.* **259**, 108160 (2022).
- Cao, B., Yang, S., Sun, A., Dong, Z. & Zhang, T. Domain knowledge-guided interpretive machine learning: formula discovery for the oxidation behavior of ferritic-martensitic steels in supercritical water. *J. Mater. Inf.* **2**, 4 (2022).
- Wei, Q. et al. Discovering a formula for the high temperature oxidation behavior of FeCrAlCoNi based high entropy alloys by domain knowledge-guided machine learning. *J. Mater. Sci. Technol.* **149**, 237–246 (2023).
- Wei, Q., Xiong, J., Sun, S. & Zhang, T. Y. Multi-objective machine learning of four mechanical properties of steels. *Sci. China Technol. Sci.* **51**, 15 (2021).
- 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.* **5**, 87 (2019).
- Jiang, L. et al. Discovery of aluminum alloys with ultra-strength and high-toughness via a property-oriented design strategy. *J. Mater. Sci. Technol.* **98**, 33–43 (2022).
- Zhang, H., Fu, H., Zhu, S., Yong, W. & Xie, J. Machine learning assisted composition effective design for precipitation strengthened copper alloys. *Acta Mater.* **215**, 117118 (2021).
- Mannodi-Kanakkithodi, A., Pilania, G., Ramprasad, R., Lookman, T. & Gubernatis, J. E. Multi-objective optimization techniques to design the Pareto front of organic dielectric polymers. *Comput. Mater. Sci.* **125**, 92–99 (2016).
- Schaefer, M., Laub, W., Sabee, J. M., Fournelle, R. A. & Lee, P. S. A numerical method for predicting intermetallic layer thickness developed during the formation of solder joints. *J. Electron. Mater.* **25**, 992–1003 (1996).
- Wu, Y. et al. The formation and growth of intermetallics in composite solder. *J. Electron. Mater.* **22**, 769–777 (1993).
- Gain, A. K., Fouzder, T., Chan, Y. C., Sharif, A. & Yung, W. K. Investigation of small Sn–3.5 Ag–0.5 Cu additions on the microstructure and properties of Sn–8Zn–3Bi solder on Au/Ni/Cu pads. *J. Alloy. Compd.* **489**, 678–684 (2010).
- Zhang, L. & Gao, L. L. Interfacial compounds growth of SnAgCu (nano La2O3)/Cu solder joints based on experiments and FEM. *J. Alloy. Compd.* **635**, 55–60 (2015).
- Pal, M. K., Gergely, G., Koncz-Horváth, D. & Gács, Z. Investigation of the electroless nickel plated sic particles in sac305 solder matrix. *Powder Metall. Met. Ceram.* **58**, 529–537 (2020).
- Gain, A. K., Chan, Y. C. & Yung, W. K. Microstructure, thermal analysis and hardness of a Sn–Ag–Cu–1 wt% nano-TiO2 composite solder on flexible ball grid array substrates. *Microelectron. Reliab.* **51**, 975–984 (2011).
- Erer, A. M. Wetting behaviors and interfacial properties of SAC300, SAC305 and SAC307 ternary Pb-Free solder alloys. *Technol. Appl. Sci.* **12**, 163–169 (2017).
- Omaç, F., Ozyurek, D. & Erer, M. Investigation of the wetting properties of ternary lead-free solder alloys on copper substrate. *Acta Phys. Pol. A* **131**, 165–167 (2017).
- Zhang, P. et al. Effect of nanoparticles addition on the microstructure and properties of lead-free solders: a review. *Appl. Sci.* **9**, 2044 (2019).
- Cheng, S., Huang, C. M. & Pecht, M. A review of lead-free solders for electronics applications. *Microelectron. Reliab.* **75**, 77–95 (2017).

35. Zaimi, N. M. et al. Effect of kaolin geopolymer ceramic addition on the properties of Sn-3.0 Ag-0.5 Cu solder joint. *Mater. Today Commun.* **25**, 101469 (2020).
36. Tsao, L. C. Evolution of nano-Ag3Sn particle formation on Cu-Sn intermetallic compounds of Sn3. 5Ag0. 5Cu composite solder/Cu during soldering. *J. Alloy. Compd.* **509**, 2326–2333 (2011).
37. Sona, M. & Prabhu, K. N. Review on microstructure evolution in Sn–Ag–Cu solders and its effect on mechanical integrity of solder joints. *J. Mater. Sci. Mater. Electron.* **24**, 3149–3169 (2013).
38. El-Daly, A. A., El-Taher, A. M. & Gouda, S. Development of new multicomponent Sn–Ag–Cu–Bi lead-free solders for low-cost commercial electronic assembly. *J. Alloy. Compd.* **627**, 268–275 (2015).
39. Kanlayasiri, K., Mongkolwongroj, M. & Ariga, T. Influence of indium addition on characteristics of Sn-0.3 Ag-0.7 Cu solder alloy. *J. Alloy. Compd.* **485**, 225–230 (2009).
40. El-Daly, A. A., El-Hosainy, H., Elmosalami, T. A. & Desoky, W. M. Microstructural modifications and properties of low-Ag-content Sn–Ag–Cu solder joints induced by Zn alloying. *J. Alloy. Compd.* **653**, 402–410 (2015).
41. Chuang, C. L., Tsao, L. C., Lin, H. K. & Feng, L. P. Effects of small amount of active Ti element additions on microstructure and property of Sn3. 5Ag0. 5Cu solder. *Mater. Sci. Eng. A* **558**, 478–484 (2012).
42. Sungkaphaitoon, P. & Plookphol, T. The effects of antimony addition on the microstructural, mechanical, and thermal properties of Sn-3.0 Ag-0.5 Cu solder alloy. *Metall. Mater. Trans. A* **49**, 652–660 (2018).
43. Sabri, M. F. M. et al. Microstructural stability of Sn-1Ag-0.5 Cu-xAl ($x=1, 1.5$, and 2 wt.%) solder alloys and the effects of high-temperature aging on their mechanical properties. *Mater. Charact.* **78**, 129–143 (2013).
44. Watanabe, H., Hidaka, N., Shohji, I. & Ito, M. Effect of Ni and Ag on interfacial reaction and microstructure of Sn-Ag-Cu-Ni-Ge lead-free solder. *Mater. Sci. Technol.* <https://citesexx.ist.psu.edu/document?repid=rep1&type=pdf&doi=654b35559d67ad489c9abc173c2896f63aa1c273> (2006).
45. Brochu, E., Cora, V. M. & De Freitas, N. A Tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. Preprint at <https://arxiv.org/abs/1012.2599> (2010).
46. Agrawal, R. Sample mean based index policies by $O(\log n)$ regret for the multi-armed bandit problem. *Adv. Appl. Probab.* **27**, 1054–1078 (1995).
47. Kushner, H. J. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *J. Basic Eng.* **86**, 97–106 (1964).
48. Jones, D. R., Schonlau, M. & Welch, W. J. Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* **13**, 455–492 (1998).
49. Bath, J. *Lead-Free Soldering*. (Springer, 2007).
50. Beáta, Š., Erika, H., & Ingrid, K. Development of SnAgCu solders with Bi and In additions and microstructural characterization of joint interface. *Weld. World* **61**, 613–621 (2017).
51. He, M., Ekpenna, S. N. & Acoff, V. L. Microstructure and creep deformation of Sn–Ag–Cu–Bi/Cu solder joints. *J. Electron. Mater.* **37**, 300–306 (2008).
52. Braga, M. H. et al. The experimental study of the Bi–Sn, Bi–Zn and Bi–Sn–Zn systems. *Calphad.* **31**, 468–478 (2007).
53. Witkin, D. Creep behavior of Bi-containing lead-free solder alloys. *J. Electron. Mater.* **41**, 190–203 (2012).
54. Moser, Z., Dutkiewicz, J., Gasior, W. & Salawa, J. The Sn–Zn (tin-zinc) system. *Bull. Alloy Phase Diagr.* **6**, 330–334 (1985).
55. El-Daly, A. A., El-Taher, A. M. & Gouda, S. Novel Bi-containing Sn-1.5 Ag-0.7 Cu lead-free solder alloy with further enhanced thermal property and strength for mobile products. *Mater. Des.* **65**, 796–805 (2015).

ACKNOWLEDGEMENTS

This work was sponsored by the National Key Research and Development Program of China (No. 2018YFB0704400), Key Program of Science and Technology of Yunnan Province (No. 202002AB080001-2), Key Research Project of Zhejiang Laboratory (No. 2021PEOAC02), Shanghai Pujiang Program (Grant No. 20PJ1403700) and Guangzhou Municipal Science and Technology Project (No. 2023A03J0003). We would like to acknowledge the support from Yunnan Tin Group (Holding) Co. Ltd, China. We also acknowledge the support from the Shanghai Engineering Research Center for Integrated Circuits and Advanced Display Materials.

AUTHOR CONTRIBUTIONS

Q.W. performed the research, conducted the experiment, analysed data, and participated in the manuscript drafting. B.C. performed the research, analysed data, encoded the algorithm of TCGPR, and participated in the manuscript drafting. H.Y. performed the research, analysed data, wrote the codes of Bayesian sampling, and participated in experiments and manuscript drafting. Y.C., K.Y., S.Y. and T.Y. analysed the experimental data. T.-Y.Z. and Z.D. supervised the project. Z.D. guided the experimental research, analysed data, and revised the manuscript. T.-Y.Z. developed the algorithm of TCGPR, guided the research, analysed data, and finalized the manuscript. All discussed the results.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41524-023-01150-0>.

Correspondence and requests for materials should be addressed to Ziqiang Dong or Tong-Yi Zhang.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, 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. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023