

RAPID COMMUNICATION

# Machine learning technique for prediction of magnetocaloric effect in $\text{La}(\text{Fe},\text{Si}/\text{Al})_{13}$ -based materials

To cite this article: Bo Zhang *et al* 2018 *Chinese Phys. B* **27** 067503

View the [article online](#) for updates and enhancements.

## Related content

- [Composition design for  \$\(\text{PrNd-La-Ce}\)\_2\text{Fe}\_{14}\text{B}\$  melt-spun magnets by machine learning technique](#)  
Rui Li, Yao Liu, Shu-Lan Zuo *et al.*
- [The magnetic properties and magnetocaloric effects in binary R-T \(R = Pr, Gd, Tb, Dy, Ho, Er, Tm; T = Ga, Ni, Co, Cu\) intermetallic compounds](#)  
Xin-Qi Zheng and Bao-Gen Shen
- [Influence of misch metal content on microstructure and magnetic properties of R-Fe-B magnets sintered by dual alloy method](#)  
Rong-Xiang Shang, Jie-Fu Xiong, Dan Liu *et al.*

# Machine learning technique for prediction of magnetocaloric effect in La(Fe,Si/Al)<sub>13</sub>-based materials\*

Bo Zhang(张博)<sup>1,2</sup>, Xin-Qi Zheng(郑新奇)<sup>3</sup>, Tong-Yun Zhao(赵同云)<sup>1,2</sup>, Feng-Xia Hu(胡凤霞)<sup>1,2</sup>,  
Ji-Rong Sun(孙继荣)<sup>1,2</sup>, and Bao-Gen Shen(沈保根)<sup>1,2,†</sup>

<sup>1</sup>State Key Laboratory of Magnetism, Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China

<sup>2</sup>School of Physical Sciences, University of the Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup>School of Materials Science and Engineering, University of Science and Technology Beijing, Beijing 100083, China

(Received 19 March 2018; revised manuscript received 4 May 2018; published online 30 May 2018)

Data-mining techniques using machine learning are powerful and efficient for materials design, possessing great potential for discovering new materials with good characteristics. Here, this technique has been used on composition design for La(Fe,Si/Al)<sub>13</sub>-based materials, which are regarded as one of the most promising magnetic refrigerants in practice. Three prediction models are built by using a machine learning algorithm called gradient boosting regression tree (GBRT) to essentially find the correlation between the Curie temperature ( $T_C$ ), maximum value of magnetic entropy change ( $(\Delta S_M)_{\max}$ ), and chemical composition, all of which yield high accuracy in the prediction of  $T_C$  and  $(\Delta S_M)_{\max}$ . The performance metric coefficient scores of determination ( $R^2$ ) for the three models are 0.96, 0.87, and 0.91. These results suggest that all of the models are well-developed predictive models on the challenging issue of generalization ability for untrained data, which can not only provide us with suggestions for real experiments but also help us gain physical insights to find proper composition for further magnetic refrigeration applications.

**Keywords:** La(Fe,Si/Al)<sub>13</sub>-based materials, composition design, machine learning, magnetic refrigeration

**PACS:** 75.30.Sg, 75.20.En

**DOI:** 10.1088/1674-1056/27/6/067503

## 1. Introduction

The magnetocaloric effect (MCE), reversible entropy or temperature variation upon the change of external magnetic fields, is an inherent quality of magnetic materials. Magnetic refrigeration (MR) based on MCE is becoming a great topic of current interest on account of the potential improvement in environmental friendliness and high efficiency compared to the conventional vapor compression methodology. MCE materials are usually characterized in terms of the maximum value of magnetic entropy change ( $(\Delta S_M)_{\max}$ ) in an isothermal process, and the Curie temperature ( $T_C$ ) around which the material works. In the last two decades, significant progress has been made in exploring large MCE materials for applications at low and room temperatures. Typical low-temperature MCE materials mainly include the rare earth-based intermetallic compounds, such as RCo<sub>2</sub>,<sup>[1,2]</sup> RNi,<sup>[3]</sup> etc. However, more attention has been paid to room temperature MCE materials because of their potential applications in refrigerators and air conditioners. These MCE materials mainly include Gd<sub>5</sub>Si<sub>2</sub>Ge<sub>2</sub>,<sup>[4,5]</sup> MnAs<sub>1-x</sub>Sb<sub>x</sub>,<sup>[6]</sup> MnFeP<sub>1-x</sub>As<sub>x</sub>,<sup>[7]</sup> Heusler alloys,<sup>[8,9]</sup> etc.

Among these MCE materials, La(Fe,Si/Al)<sub>13</sub>-based compounds are commonly approved as one of the most promising

magnetic refrigerants. The binary LaFe<sub>13</sub> phase does not exist as a result of its positive formation enthalpy.<sup>[10]</sup> Therefore, a third element, either Si or Al, had to be introduced to stabilize a cubic NaZn<sub>13</sub>-type structure. For La(Fe,Si/Al)<sub>13</sub>-based compounds, there are many studies focusing on improving its MCE performance by, for example, changing the Fe and Si/Al content,<sup>[10,11]</sup> doped transitional metal (Co and Mn),<sup>[12,13]</sup> rare earth (Ce, Pr, Nd),<sup>[14–16]</sup> interstitials (H, C, B),<sup>[17–19]</sup> etc. When the chemical composition of La(Fe,Si/Al)<sub>13</sub>-based compounds is changed,  $(\Delta S_M)_{\max}$  and  $T_C$  may also change. Predicting these physical features is not easy when the composition is complex. However, data-mining techniques using machine learning can be applied to build models for the prediction of the relationship between physical features with composition-related parameters that can be obtained prior to experiment.

Machine learning is a field of computer science that provides computer systems with the ability to “learn” (i.e., progressively improve performance on a specific task) from data without being explicitly programmed. Machine learning is widely used in many fields after decades of development. It can build a predictive model between features and target property by using statistical rules from massive experimental data. It is a powerful method to analyze large volume and multidimensional

\*Project supported by the National Basic Research Program of China (Grant No. 2014CB643702), the National Natural Science Foundation of China (Grant No. 51590880), the Knowledge Innovation Project of the Chinese Academy of Sciences (Grant No. KJZD-EW-M05), and the National Key Research and Development Program of China (Grant No. 2016YFB0700903).

†Corresponding author. E-mail: shenbg@iphy.ac.cn

mensional data that are perfectly matched for materials design. Materials design that involves machine learning has been used successfully in various research fields.<sup>[20–24]</sup> For example, Anton *et al.*<sup>[22]</sup> built a machine learning model to make fast and reliable predictions of the occurrence of Heusler versus non-Heusler compounds for an arbitrary combination of elements with no structural input on over  $4 \times 10^5$  candidates. Xue *et al.*<sup>[23]</sup> built a model between composition and thermal hysteresis ( $\Delta T$ ) in  $\text{Ti}_{50}\text{Ni}_{50-x-y-z}\text{Cu}_x\text{Fe}_y\text{Pd}_z$  shape memory alloys and ultimately discovered 14 new compounds possessing smaller  $\Delta T$  than the original dataset from a potential space of  $\sim 8 \times 10^5$  compositions. Similar methods were also used in the MCE field. Taoreed *et al.*<sup>[24]</sup> built a model to predict the Curie temperature ( $T_C$ ) of manganite-based materials in different dopants that found a proper manganite quickly.

To date, data-mining techniques using machine learning have not been reported in the materials design of  $\text{La}(\text{Fe},\text{Si}/\text{Al})_{13}$ -based compounds. In this paper, we successfully developed prediction models using a machine learning algorithm called gradient boosting regression tree (GBRT)<sup>[25]</sup> to predict the maximum value of magnetic entropy change ( $(\Delta S_M)_{\max}$ ) and the Curie temperature ( $T_C$ ) of  $\text{La}(\text{Fe},\text{Si}/\text{Al})_{13}$ -based materials based on chemical composition, which have the potential to boost the speed of composition design research fundamentally and provide implications for solving the challenging and elusive issue of finding a proper composition for further application.

## 2. Data collection and model setup

The dataset of  $\text{La}(\text{Fe},\text{Si}/\text{Al})_{13}$ -based materials was collected from the published literatures and contained 144 pieces of data.<sup>[26]</sup> Noting that the divergences of phase, Curie temperature, and  $(\Delta S_M)_{\max}$  for materials may reduce the accuracy of models' prediction, all of the data we collected here

are guaranteed to be  $\text{NaZn}_{13}$ -type structure, and the values of Curie temperatures and  $(\Delta S_M)_{\max}$  are reliable for model learning. The data feature we used here is chemical composition, and the target properties we predicted are Curie temperature  $T_C$  and  $(\Delta S_M)_{\max}$  under a magnetic field change of 0 T–2 T and 0 T–5 T, respectively. Because some data from the dataset may not contain all three of the target properties, the dataset we used to build the prediction model for each target property is not the same. For each target, we used 141 data points for  $T_C$ , 108 for  $(\Delta S_M)_{\max}$  at 0 T–2 T, and 93 for  $(\Delta S_M)_{\max}$  at 0 T–5 T.

The task we faced in supervised learning of traditional machine learning field is a regression problem. The algorithms that can handle this task mainly contain linear regression, decision tree regression, support vector regression, and some ensemble methods. The algorithm we used here is GBRT, which belongs to the ensemble methods. The GBRT algorithm produces a prediction model in the form of an ensemble of weak prediction models, which is an ensemble of regression tree models. The algorithm trains weak regression tree models once to make a weak prediction each time. Although the prediction errors are large at first, the algorithm will iterate training new weak regression trees to make the prediction better. Each new tree helps to correct errors made by the previously trained trees. Boosting is based on the idea of whether a weak model can be modified to become better. There are two important parameters in the GBRT algorithm: one is the number of estimators, which means the number of weak regression tree models; the other is the learning rate, which means the speed of improvement during iterating training weak regression trees. The GBRT algorithm is considered to be one of the best methods in machine learning. It can fit data where the relationship between features and target properties is complex well and perform robustly when facing outliers.

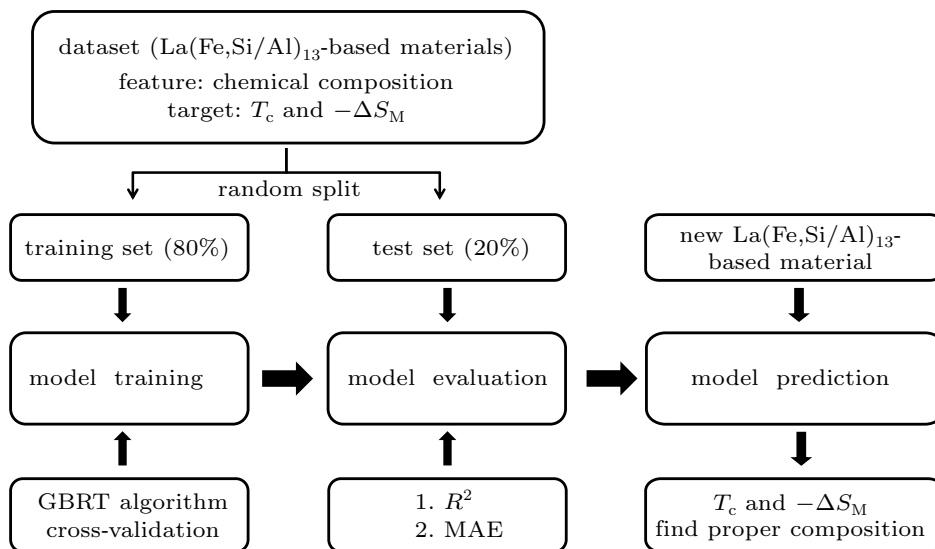


Fig. 1. Overall setup processes for prediction models by using machine learning approach on  $\text{La}(\text{Fe},\text{Si}/\text{Al})_{13}$ -based materials.

From the dataset, we used the chemical composition as a data feature and Curie temperature ( $T_C$ ) and  $(\Delta S_M)_{\max}$  as target properties. Our models will learn from the data to predict target properties via some features. The overall setup for prediction models is shown in Fig. 1. There are three main processes:

(i) Split the dataset into two parts: training set and test set. The training set is used to train the model, and the test set is used to evaluate the performance of the trained model. In general, the training set is 80% of the entire dataset, and the test set is 20%. To optimize performance, the dataset must be split randomly.

(ii) Build the GBRT model. Then, search for the best parameters of the model by the grid-search method.<sup>[27]</sup> The performance metric is  $R^2$  for the model prediction of training set after using  $k$ -fold cross-validation.  $R^2$  represents the variance explained by the model (the higher, the better), and its maximal value is equal to 1.

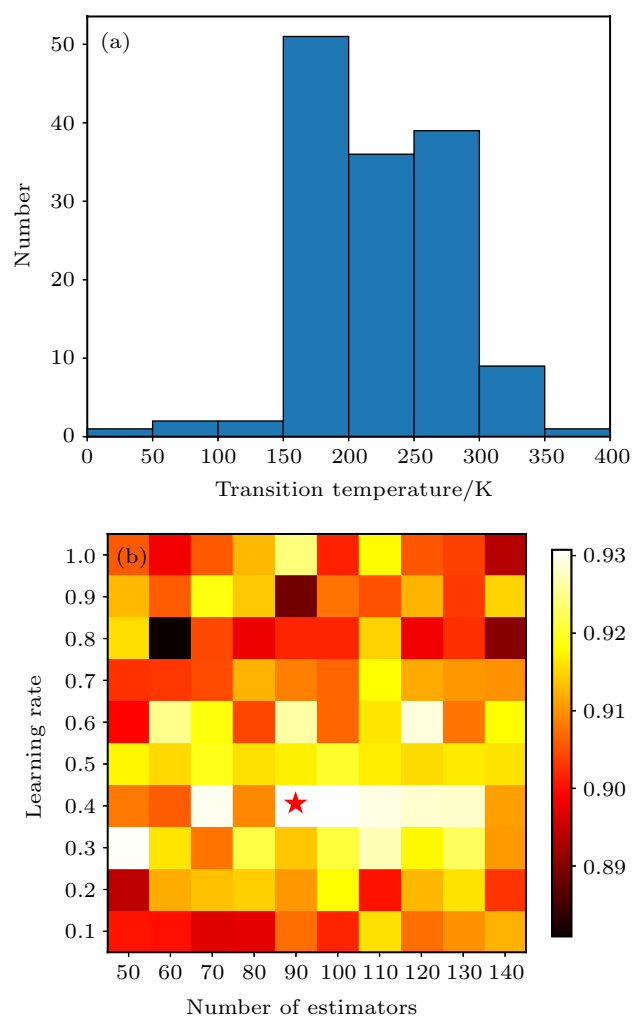
(iii) Evaluate the model's predictive performances on the test set, as soon as model training finished. The metrics are  $R^2$  and MAE (mean absolute error: the lower, the better).

Finally, the model can be used to estimate target properties for new materials via inputs of chemical composition, which can serve as a quick guide to composition design and help find proper materials.

In this paper, all of the models' setup and analysis are implemented in Python 3.6 with the scikit-learn open-source package.

### 3. Results and discussion

We used 108 data points to train the  $T_C$  prediction model. The data distribution is counted. As shown in Fig. 2(a), the Curie temperatures  $T_C$  of the data are all distributed in the range of 0 K–400 K. Among those, the majority scatters in the range of 150 K–300 K, covering the room temperature. We randomly split the dataset between a training set (80%) and a test set (20%) and used the training set to train the model we built. As the training set is relatively small, fivefold cross-validation is used on the training set data. In order to find the best values of parameters (number of estimators and learning rate) for the model, we use grid-search, which combines the number of estimators and the learning rate with different values to train the model, and find the parameters that make up the best  $R^2$  for the model as the best prediction model. For this procedure, we chose learning rate from 0.1–1 and number of estimators from 50–140, respectively, for all models. The results are depicted in Fig. 2(b). We trained 100 combinations of two parameters models and found that the best combination is 90 for number of estimators and 0.4 for learning rate, which is regarded as our best model.



**Fig. 2.** (color online) (a) Data distribution for  $T_C$  dataset. (b) Performances of  $T_C$  prediction models trained with different parameter combinations. The best parameter combination is marked with red star.

After the best model is trained, it needs to be evaluated to determine whether it is feasible. A general way is to use the test set for the untrained data. The metrics we used to evaluate are  $R^2$  and MAE. The results are depicted in Fig. 3. The  $X$  axis is the predicted Curie temperature from the model, and the  $Y$  axis is the measured one. If the model is perfect, the predicted  $T_C$  will be exactly the same as the measured one, and all data points will align along the  $45^\circ$  diagonal line. The result of training set is shown in Fig. 3(a), which shows great fitting on measured  $T_C$ . The result with 0.98 of  $R^2$  and 2.03 K of MAE shows the model had been trained perfectly on the training set. The result of evaluation of the model on the test set is shown in Fig. 3(b).  $R^2$  is 0.96, and the MAE is 9.81 K, which indicates the model has good generalization ability for untrained data. In addition, that the training result is slightly higher compared to the test result, as shown in Fig. 3, implies that the model is a little overfit, which means the model prediction corresponds too closely or exactly for the training set and may result in failing to fit additional data reliably. However, because the number of datasets is relatively small, model overfitting is unavoidable in principle. Overall, the  $T_C$  prediction model corresponds well with the generalization.

The number of data points used to train  $(\Delta S_M)_{\max}$  prediction models for 0 T–2 T and 0 T–5 T is 108 and 93, respectively. The data distributions are shown in Figs. 4(a) and 4(c).  $(\Delta S_M)_{\max}$  at 0 T–2 T is evenly distributed from 0 to 30 J·kg<sup>-1</sup>·K<sup>-1</sup>, with a peak in the bin of 5 J·kg<sup>-1</sup>·K<sup>-1</sup>–10 J·kg<sup>-1</sup>·K<sup>-1</sup>.  $(\Delta S_M)_{\max}$  at 0 T–5 T is evenly distributed from 5 J·kg<sup>-1</sup>·K<sup>-1</sup> to 30 J·kg<sup>-1</sup>·K<sup>-1</sup>, with maximum in the

bin of 25 J·kg<sup>-1</sup>·K<sup>-1</sup>–30 J·kg<sup>-1</sup>·K<sup>-1</sup>. The process of training and selecting parameters is almost the same as that of  $T_C$ . The results for 0 T–2 T and 0 T–5 T are shown in Figs. 4(b) and 4(d), respectively. The best combinations are (100, 0.2) for  $(\Delta S_M)_{\max}$  at 0 T–2 T, and (80, 0.1) for  $(\Delta S_M)_{\max}$  at 0 T–5 T, respectively. We used these combinations as the best models.

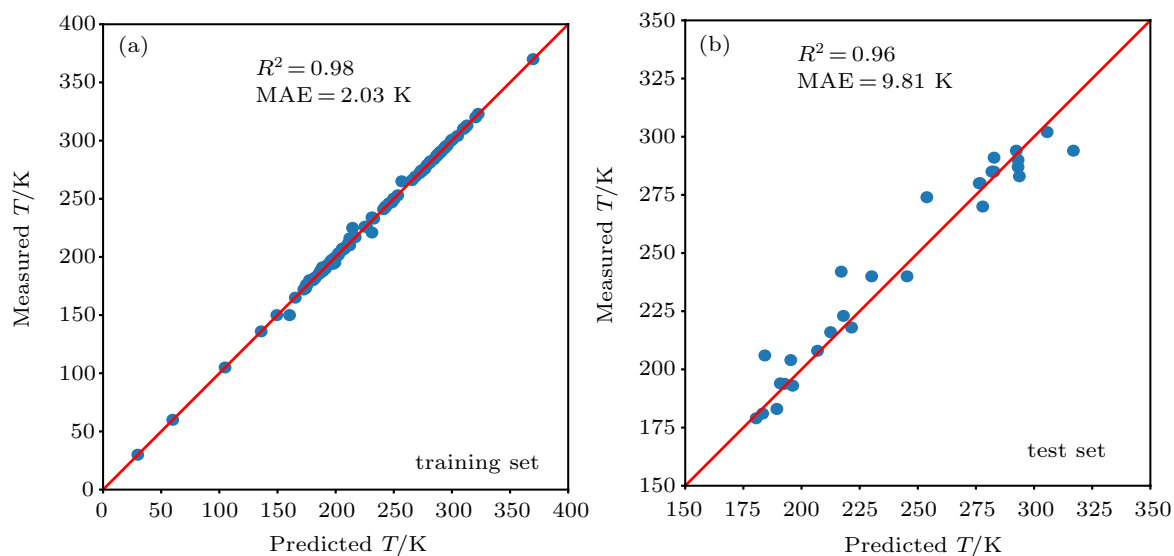


Fig. 3. (color online) Predicted scatter plots on  $T_C$  dataset. (a) Training set and (b) test set.

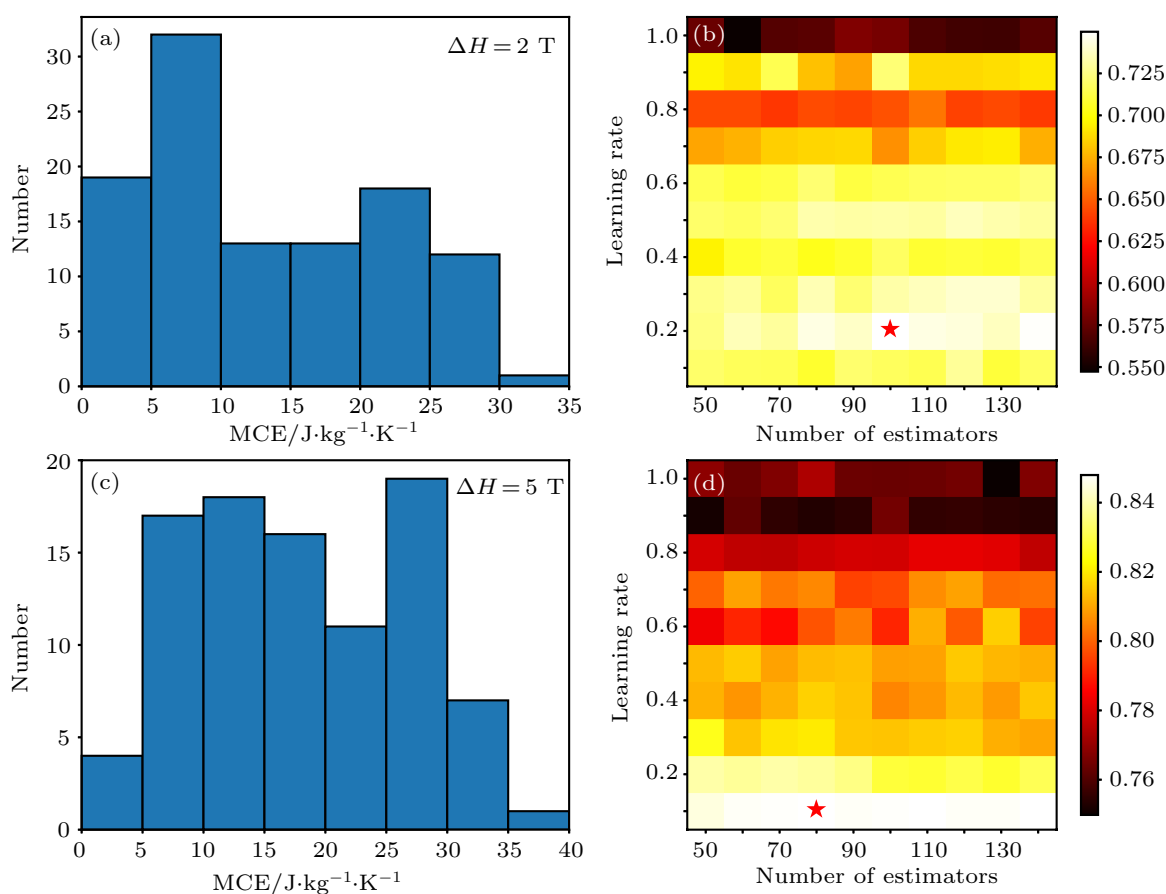


Fig. 4. (color online) (a) and (c) Data distribution for  $(\Delta S_M)_{\max}$  at 0 T–2 T and 0 T–5 T datasets, respectively. (b) and (d) Performances of  $(\Delta S_M)_{\max}$  at 0 T–2 T and 0 T–5 T prediction models trained with different parameters combinations, respectively. The best parameters combinations are marked with red star.

After training the best models, we use the test set to evaluate the models. The metrics and process are the same as for the  $T_C$  prediction model. The two models' results are shown in Fig. 5. The results of the training set with  $R^2 = 0.98$  and  $MAE = 0.60 \text{ J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$  for 0 T–2 T and  $R^2 = 0.97$  and  $MAE = 1.14 \text{ J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$  for 0 T–5 T, respectively, are shown in Figs. 5(a) and 5(c), which indicate the models were well trained with the training set. The results

of evaluation of the model for the test set with  $R^2 = 0.87$  and  $MAE = 2.51 \text{ J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$  for 0 T–2 T and  $R^2 = 0.91$  and  $MAE = 2.53 \text{ J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$  for 0 T–5 T, respectively, are obtained and shown in Figs. 5(b) and 5(d). All the models have good generalization ability. The results also show the two models are somewhat overfit for the same reason as the  $T_C$  prediction model, which is unavoidable.

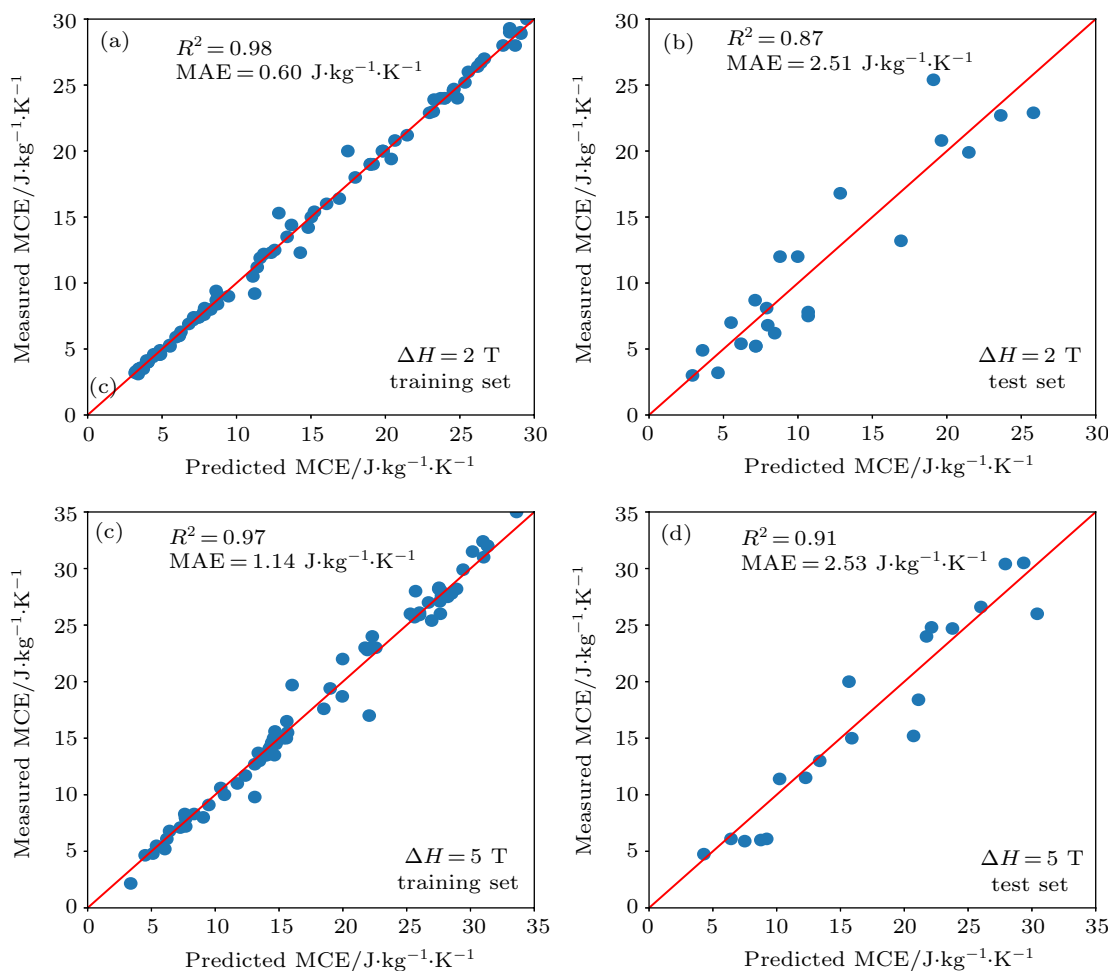


Fig. 5. (color online) Predicted scatter plots on  $(\Delta S_M)_{\max}$  at 0 T–2 T and 0 T–5 T datasets. (a) and (c) Training set, (b) and (d) test set.

All of the models we built above have good predictive performance. The model with  $(\Delta S_M)_{\max}$  at 0 T–2 T performs similarly to that with  $(\Delta S_M)_{\max}$  at 0 T–5 T. Compared to the  $(\Delta S_M)_{\max}$  prediction models, the  $T_C$  prediction model has better performance. Because the datasets of the three models possess similar size, the phenomenon may result from the magnetic properties that the models predict. The Curie temperature,  $T_C$ , is an intrinsic magnetic property for materials that is hardly affected by the measurement methods. By contrast,  $(\Delta S_M)_{\max}$  is a calculated value from experimental results, and is sensitive to the detailed measurement methods and experiments. Considering the complexity of obtaining  $(\Delta S_M)_{\max}$ , the datasets for  $(\Delta S_M)_{\max}$  may be of larger errors, which may give rise to more uncertainty to predicate  $(\Delta S_M)_{\max}$  than  $T_C$ .

#### 4. Conclusion

In summary, we successfully developed three prediction models using a machine learning algorithm called GBRT to predict two important magnetic properties of MCE materials:  $(\Delta S_M)_{\max}$  and  $T_C$ , respectively, for  $\text{La}(\text{Fe},\text{Si}/\text{Al})_{13}$ -based materials based on the chemical composition. Based on this model, the properties of untrained data can be well predicted. Our findings suggest that machine learning is very powerful and efficient tools which can be used to accelerate composition design and find proper composition for further applications in practice.

In addition, the selection and construction of features are the most important factors affecting the performance of the

models. In further research, additional physical properties, such as lattice constant and saturation magnetization, should be included in the features as they can greatly affect the target properties. Constructing more physical features for models may significantly improve the performance.

## References

- [1] Duc N H and Anh D T K 2002 *J. Magn. Magn. Mater.* **242** 873
- [2] Giguere A, Foldeaki M, Schnelle W and Gmelin E 1999 *J. Phys.: Condens. Matter* **11** 6969
- [3] Zheng X Q, Zhang B, Wu H, Hu F X, Huang Q Z and Shen B G 2016 *J. Appl. Phys.* **120** 163907
- [4] Pecharsky V K and Gschneidner K A 1997 *Phys. Rev. Lett.* **78** 4494
- [5] Giguère A, Foldeaki M, Ravi G B, Chahine R, Bose T K, Frydman A and Barclay J A 1999 *Phys. Rev. Lett.* **83** 2262
- [6] Wada H and Tanabe Y 2001 *Appl. Phys. Lett.* **79** 3302
- [7] Tegus O, Brück E, Buschow K H J and de Boer F R 2002 *Nature* **415** 150
- [8] Hu F X, Shen B G, Sun J R and Wu G H 2001 *Phys. Rev. B* **64** 132412
- [9] Liu J, Gottschall T, Skokov K P, Moore J D and Gutfleisch O 2012 *Nat. Mater.* **11** 620
- [10] Shen B G, Sun J R, Hu F X, Zhang H W and Cheng Z H 2009 *Adv. Mater.* **21** 4545
- [11] Hu F X, Shen B G, Sun J R and Zhang X X 2000 *Chin. Phys.* **9** 550
- [12] Balli M, Fruchart D and Gignoux D 2007 *J. Phys.: Condens. Matter* **19** 236230
- [13] Barcza A, Katter M, Zellmann V, Russek S, Jacobs S and Zimm C 2011 *IEEE Trans. Magn.* **47** 3391
- [14] Fujieda S, Fujita A and Fukamichi K 2005 *IEEE Trans. Magn.* **41** 2787
- [15] Fujieda S, Fujita A and Fukamichi K 2007 *J. Appl. Phys.* **102** 023907
- [16] Shen J, Li Y X, Zhang J, Gao B, Hu F X, Zhang H W, Chen Y Z, Rong C B and Sun J R 2008 *J. Appl. Phys.* **103** 07B317
- [17] Fujita A, Fujieda S, Hasegawa Y and Fukamichi K 2003 *Phys. Rev. B* **67** 104416
- [18] Zhang H, Shen J, Xu Z Y, Zheng X Q, Hu F X, Sun J R and Shen B G 2012 *J. Magn. Magn. Mater.* **324** 484
- [19] Chen X A, Chen Y G and Tang Y B 2011 *J. Alloys Compd.* **509** 2864
- [20] Nosengo N 2016 *Nature* **533** 23
- [21] Meredig B, Agrawal A, Kirklin S, Saal J E, Doak J W, Thompson A, Zhang K, Choudhary A and Wolverton C 2014 *Phys. Rev. B* **89** 094429
- [22] Oliynyk A O, Antono E, Sparks T D, Ghadbeigi L, Gaultois M W, Meredig B and Mar A 2016 *Chem. Mater.* **28** 7324
- [23] Xue D, Balachandran P V, Hogden J, Theiler J, Xue D and Lookman T 2016 *Nat. Commun.* **7** 11241
- [24] Owolabi T O, Akande K O, Olatunji S O, Alqahtani A and Aldhafferi N 2016 *AIP Adv.* **6** 105009
- [25] Friedman J H 2001 *Ann. Stat.* **29** 1189
- [26] Franco V, Blázquez J S, Iplus J J, Law J Y, Moreno-Ramírez L M and Conde A 2018 *Prog. Mater. Sci.* **93** 112
- [27] Sun Y T, Bai H Y, Li M Z and Wang W H 2017 *J. Phys. Chem. Lett.* **8** 3434