ARTICLE OPEN

Check for updates

Extracting local nucleation fields in permanent magnets using machine learning

Markus Gusenbauer [1]✉, Harald Oezelt [1], Johann Fischbacher [1], Alexander Kovacs[1], Panpan Zhao[2], Thomas George Woodcock [2] and Thomas Schrefl[1]

Microstructural features play an important role in the quality of permanent magnets. The coercivity is greatly influenced by crystallographic defects, like twin boundaries, as is well known for MnAl-C. It would be very useful to be able to predict the macroscopic coercivity from microstructure imaging. Although this is not possible now, in the present work we examine a related question, namely the prediction of simulated nucleation fields of a quasi-three-dimensional (rescaled and extruded) system constructed from a two-dimensional image. We extract features of the image and analyze them via machine learning. A large number of extruded systems are constructed from 10×10 pixel sub-images of an Electron Backscatter Diffraction (EBSD) image using an automated meshing procedure. A local nucleation field is calculated by micromagnetic simulation of each quasi-three-dimensional system. Decision trees, trained with the simulation results, can predict nucleation fields of these quasi-three-dimensional systems from new images within seconds. As for now we cannot quantitatively predict the macroscopic coercivity, nevertheless we can identify weak spots in the magnet and see trends in the nucleation field distribution.

INTRODUCTION

Permanent magnets are of great interest in today's economy. Especially green energy applications, such as in wind turbines and hybrid/electric vehicles demand high-performance permanent magnets. The performance of permanent magnets is mainly determined by intrinsic magnetic properties and microstructural features. The intrinsic properties are adjusted by including rare earth (RE) elements, which mainly increase the magnetocrystalline anisotropy. Due to economical and environmental considerations, there is a high interest in reducing the use of RE elements[1]. MnAl-C contains no RE or other critical raw materials and its intrinsic magnetic properties make it an attractive alternative to certain types of RE-based permanent magnets[2]. The microstructure of MnAl-C magnets is known to contain a range of defects, such as grain boundaries, twins, and antiphase boundaries[3–7]. Kronmüller and Goll discussed microstructural properties of advanced hard magnetic materials and its relation to the quality of permanent magnets[8]. By investigating the various boundary types of MnAl-C, their distribution and the effect on the coercivity by micromagnetic simulations and machine learning, we hope to optimize the overall performance of MnAl-C magnets.

Micromagnetic simulations can be a vital tool to improve the understanding of permanent magnets and their macroscopic properties. Widely used measures to determine the quality of permanent magnets are the nucleation field H_n, the coercive field H_c and the energy density product BH_{max}[9]. Here we denote the critical value of the external field at which the magnetization starts to switch irreversibly as nucleation field H_n[10]. First thing to know is the microstructure of the magnetic material. In a second step an accurate theory is needed to determine the quality measures from this microstructure. Micromagnetic simulations can calculate demagnetization processes and further on H_n, H_c, and BH_{max} with the use of the present microstructure and the intrinsic magnetic properties of the material. Intrinsic magnetic properties can be

obtained either from ab initio simulations or from the measurements of the anisotropy constant, the magnetization and the exchange constant. In addition to the intrinsic magnetic properties the microstructure of permanent magnets influences magnetization reversal. A useful technique to get local crystallographic orientations from a microstructure is Electron Backscatter Diffraction (EBSD) microscopy[11]. It allows to analyze large areas in comparison to transmission electron microscopy (TEM), with the disadvantage of a worse resolution. The smallest spatial resolution of standard EBSD is about 30–50 nm. Creating an EBSD image with about half a million pixels takes a few hours on modern systems.

In order to create a micromagnetic simulation from an EBSD image, several steps are necessary. The EBSD raw data need to be scanned, the microstructure needs to be extracted and the underlying crystallographic orientation needs to be stored. The microstructure is then meshed with very small finite elements (tetrahedrons). Only now a micromagnetic calculation can be performed. The accuracy of micromagnetic simulations with a known microstructure of a magnet is well described, as for example in ref. [12]. The simulations nicely reproduce trends, yet the absolute value of the simulation results is always given with a certain offset to the experiments. The simulations cannot incorporate all effects, as for example inaccurate or unknown phase distributions. The size limitation in the simulations affects coercivity as well, because with larger samples (larger grains), the coercivity is decreasing[13]. The same offset from experiments to the simulations has been observed for MnAl-C[6,14]. In this work we are interested in trends of the nucleation field distribution and to find weak spots in the microstructure, which can still be done in spite of this offset.

The size of a micromagnetic simulation model is limited. The intrinsic length to represent magnetization reversal, the so-called characteristic length scale, need to be chosen such, that nucleation processes or domain wall movements can be reflected

[1]Department for Integrated Sensor Systems, Danube University Krems, Krems an der Donau, Austria. [2]Institute for Metallic Materials, Leibniz IFW Dresden, Dresden, Germany. ✉email: markus.gusenbauer@donau-uni.ac.at

within the finite element mesh. The largest allowed length of a finite element is typically in the range of a few nm only. There are two ways to calculate the demagnetization curve numerically with micromagnetic solvers:

(1) With state-of-the-art micromagnetic energy minimization codes it is possible to calculate the demagnetization curve of a permanent magnet with a size of $250 \times 250 \times 250 \text{ nm}^3$ within 1–2 days[15]. Using a finite differences code, optimized for massively parallel computing on a supercomputer, a cube with the size of about $1 \times 1 \times 1 \text{ μm}$ can be solved within a reasonable time. For example, using 256 CPUs the authors report a simulation time of 6.68 days for computing a demagnetization curve[16].

(2) Solving time-dependent Landau–Lifshitz–Gilbert (LLG) equations reduces the practical model size even further. With the software applied in this work[17] the computation of the demagnetization curve of a cube with a size of $200 \times 200 \times 200 \text{ nm}^3$ on a single CPU lasts about 3–6 days depending on the magnetic configuration.

Alternatively to the simulation of single large systems, several small parts, each covering a specific aspect of the microstructure, can be simulated. However, computing many small samples can be challenging again in terms of available computational resources. Predicting the coercive field directly from a microstructural image would speed up the computation tremendously. A promising approach to reduce the computational costs is the use of machine learning to predict the coercivity of permanent magnets[18]. Even if it is only a rough approximation, one can gain qualitative information of the influence of local microstructural properties on the coercivity. In material sciences the use of machine learning has helped determine microstructure-properties relationships, automates and increases the speed in materials characterization or fosters materials discovery[19]. Size limitations can be overcome by e.g., bridging the gap between molecular dynamics and macroscopic materiel sciences[20]. Machine learning has been used for the characterization of e.g., steel microstructures[21]. In Mg-alloys, Orme et al. analyzed the formation and propagation of twinning boundaries using decision trees[22].

Typically in machine learning a large amount of data is required. After acquisition, the data need to be prepared to fit a chosen machine learning model that is suitable for the specific problem. The model is fed with the data to incrementally improve the predictive ability. Evaluation is performed on parts of the data, that have not been used for training. Finally, a completely new set of input data is tested. The errors are computed for this test set, which is data the model have not seen before. In this work we obtain the training data directly from EBSD data of MnAl-C and micromagnetic simulations[23]. Using automated geometry construction and discretization of the EBSD map, we can compute local nucleation fields for many small image extracts (selections) using micromagnetic simulations. Thus a large amount of input data can be acquired. We tried random forest and gradient boosting regressors[24] for learning local nucleation fields H_n. We also show that, when fitting H_n to a set of selected properties derived from the EBSD image, a simple linear regression model gives good results. The models were evaluated and their hyperparameters optimized by averaging the error of a 5-fold cross-validation. Afterwards we test the prediction performance of the initial training data on a completely new EBSD map. Generating the necessary data for a machine learning model once takes very long (weeks or months), yet training and predicting then only takes a few minutes. Even though the prediction contains errors, the quality of a magnet can be analyzed and good or bad spots in the magnet's microstructure can be found.

In other fields of material science it is common practice to use machine learning for characterizing the influence of microstructure on materials properties[19]. This work is a first attempt to bring machine learning to the magnetism community. We think our method is useful from several points of view:

(1) Machine learning is a building block for multiscale simulations.
(2) Machine learning helps to identify weak spots with local nucleation fields in the microstructure.
(3) Machine learning can contribute and act as an additional tool that complements magnetic measurements and micromagnetic simulations.

RESULTS

In the Results section we show feature importance of our machine learning model using different approaches:

(1) We predict nucleation fields for different feature vectors (single features or feature combinations).
(2) Partial dependence plots show the relation from features to predicted results.
(3) A linear regression model highlights the strength of important derived features.

Further on we discuss problems with complex structures such as grain boundary junctions. And we test the trained machine on a new EBSD dataset (EBSD-B) to demonstrate the predictive ability. In the end we suggest a way to use the predicted nucleation fields for obtaining bulk magnetic properties.

Feature importance by feature vector

In section Feature engineering we defined five features from our microstructural EBSD dataset. Not all features might be equally important to accurately predict the nucleation field. In order to see the feature importance we predicted the local nucleation field via a single feature and combinations of features. Table 1 shows the mean absolute error (MAE) and root mean squared error (RMSE) for the various models trained with different feature vectors. Please note that the hyperparameters of the regressors were optimized just once using all features. Afterwards they are applied to all feature combinations. A separate optimization for each feature combination gave no significant improvement of the prediction performance.

We compare predictions for the same map (EBSD-A) as well as for a new map (EBSD-B) as outlined in Fig. 1a, b, respectively. The lowest MAE is reached with the feature GRAIN, in which the location information of the grains is not included. The PIXEL feature, which contains the additional position information, performs a little worse. Probably redundant data and the high number of features (large feature vector **F**) make the training more difficult[25]. The feature TWIN produces still an acceptable score, whereas, SW and SIZE should only be used in addition to other features.

The combination of several features introduces more possibilities for accurately predicting the nucleation field. The best feature combinations are shown in Table 1. Although a combination of many features should increase the prediction accuracy, the best performance is given with the features GRAIN, SIZE, SW, and TWIN. Due to the dependency of PIXEL and GRAIN, which introduces a lot of redundant information to the feature vector **F**[25], the prediction accuracy is not improving, when adding the PIXEL feature.

In Fig. 2a we show predictions for the local nucleation fields in EBSD-A calculated by our voting regressors using all five features (PIXEL, GRAIN, SIZE, SW, and TWIN). Local nucleation fields are nicely reconstructed with lower values at grain boundaries and higher values for bulk as shown in the simulations (Fig. 3c). For visual comparison to the previous image we are using only GRAIN

for predicting the local coercivity of the same EBSD map (Fig. 2b). Although the differences in MAE and RMSE are quite large in Table 1, the visual comparison shows a very similar trend. This trend is also found in simplified hysteresis curves constructed from the predicted nucleation fields in section Bulk magnetic properties.

Feature importance by partial dependence plots

A partial dependence plot highlights the effect of a given feature, while averaging the contributions of the other features[26]. In Fig. 4a we show one-way partial dependence plots of the features SW and TWIN, whereby the model was trained with a feature vector containing all features: PIXEL, GRAIN, SIZE, SW, and TWIN. We know from literature that SW is important, clearly the SW dependence plot shows an almost linear relation to the predicted values. For the TWIN feature we observe a similar trend, yet with a

much stronger influence on the predictions. Large misalignment between adjacent grain's orientations cause a drastic reduction in H_n. A two-way partial dependence curve of SW and TWIN in Fig. 4b shows, that best predicted nucleation fields can be expected with a large feature value of both, SW and TWIN. Here again the strong influence of the TWIN feature is visible, strong misalignment drastically reduces H_n even for a large SW value. Partial dependence plots for PIXEL, GRAIN, and SIZE are not reasonable, as each one of it contains multiple features (see section Feature engineering).

Feature importance by linear regression

Based on a qualitative understanding of magnetization reversal a simple model for nucleation fields would be a linear regressor having SW and TWIN as features. The importance of these two features is also shown in Fig. 4. Therefore we trained a linear regressor with an ordinary least square fit. We applied a 5-fold cross-validation to see the performance of the model on EBSD-A (Fig. 1a) and we tested the model on the new map EBSD-B (Fig. 1b). In Table 2 we show the MAE and RMSE scores of the linear regression model with the features SW and TWIN as input feature vector. The mean cross-validation scores are slightly worse than that obtained for the optimal feature combination with the voting regressor (see Table 1). However, when tested with previously unseen data, the linear regressor provides results comparable to the voting regressor. The MAE and RMSE on the test set (EBSD-B in Fig. 5) were 0.166 T/μ_0 and 0.245 T/μ_0 with the feature combination SW and TWIN. Values for predicting the new EBSD map are close to those of the voting regressor, which uses GRAIN, SIZE, SW, and TWIN as features.

The comparison of the results for the optimized voting regressor and the linear model shows that the voting regressor fits better to the training data. Nevertheless both models achieve similar performance measures for previously unseen data. The very good results from the linear regressor suggests as well, that the features SW and TWIN are very powerful in the proposed machine learning model.

Microstructural complexity

Right now we have used all the simulation data of EBSD-A (Fig. 3) for training and testing. For better understanding of the MAE and RMSE values, we can split the data according to the number of grains in the selections. Each dataset of EBSD-A with the same number of grains is individually used for training and testing (80% training data, 20% test data). With increasing number of grains, the MAE of the prediction is increasing as well, which is shown in Table 3. A large number of grains in the selections increases the systems complexity with many possibilities of magnetization reversal. Here most likely the number of input data has to be extended, in order to improve the prediction accuracy. The RMSE value is rather high with more than 1 grain, which indicates again the lack of data.

Table 1. Machine learning scores.

EBSD-A		EBSD-B		Active features				
MAE	RMSE	MAE	RMSE	PI	GR	SI	SW	TW
0.188	0.266	0.220	0.290	×				
0.175	0.267	0.206	0.284		×			
0.263	0.341	0.278	0.367			×		
0.242	0.337	0.328	0.426				×	
0.200	0.284	0.222	0.321					×
0.167	0.258	0.196	0.281				×	×
0.146	0.236	0.170	0.242		×			×
0.146	0.220	0.175	0.245	×	×			×
0.149	0.222	0.178	0.248	×		×		×
0.142	0.221	0.167	0.235	×			×	×
0.135	0.218	0.161	0.228		×	×		×
0.144	0.236	0.163	0.234		×		×	×
0.142	0.216	0.174	0.243	×	×	×		×
0.138	0.217	0.169	0.238	×	×		×	×
0.137	0.216	0.166	0.234	×		×	×	×
0.132	0.218	**0.157**	**0.223**		×	×	×	×
0.134	**0.213**	0.164	0.233	×	×	×	×	×

Mean absolute error (MAE [T/μ_0]) and root mean squared error (RMSE [T/μ_0]) from the voting regressor (VR) are given for single feature predictions and best feature combinations. The multiplication sign (×) marks the activated feature(s). For each scoring column the number in bold font indicates the best performance. Results for EBSD-A (Fig. 3) are obtained from the average of a 5-fold cross-validation (Fig. 1a). The results for EBSD-B (Fig. 5) shows the performance of the models on the test set after training the model with selections from EBSD-A (Fig. 1b). All feature combinations as well as the values for random forest (RF) and gradient boosting (GB) regressor can be found in Supplementary Table 1.
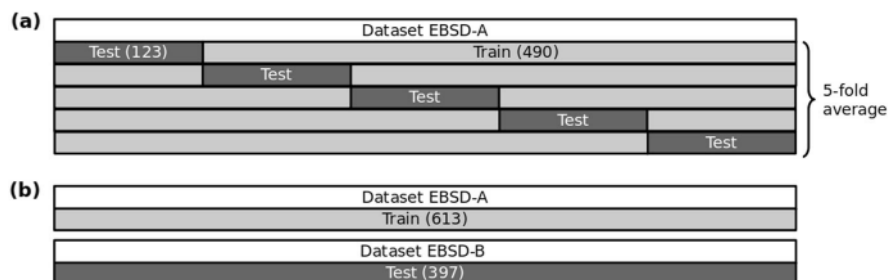

Fig. 1 Training and testing. a A 5-fold cross-validation is performed for the dataset EBSD-A (Fig. 3). **b** The full dataset EBSD-A is used for predictions in a new map EBSD-B (Fig. 5).
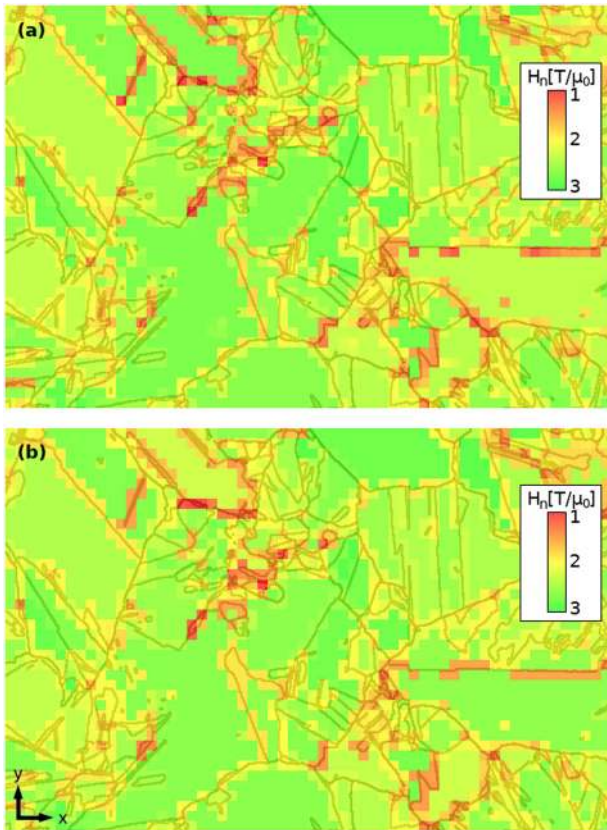
Fig. 2 EBSD-A prediction maps. Nucleation fields H_n predicted for EBSD-A (Fig. 3) by using **a** PIXEL, GRAIN, SIZE, SW, TWIN, and **b** GRAIN as training features.

Fig. 3 Training map EBSD-A. a, b EBSD map of 600×400 pixels with a pixel edge length of 0.3 µm divided in two halves. **a** The left half of the figure shows an inverse pole figure (IPF) map converted with Dream3D (dream3d.bluequartz.net, last visited on 17/12/2019). **b** The right half of the figure shows crystallographic twin boundaries: true-twins (red), order-fault-twins (green), pseudo-twins (blue), and others (gray). **c** The same map is used to depict the boundaries and the micromagnetically computed nucleation fields H_n. Each square corresponds to a unique selection (10×10 pixels) color-coded with H_n. Red squares show weak spots in the microstructure.

Prediction accuracy of new EBSD maps

With the already trained and validated machine learning model from EBSD-A (Fig. 3), any new EBSD map from the same material can be predicted in a matter of seconds rather than weeks as needed for micromagnetic simulations. To determine the prediction accuracy for new maps we test the model trained by EBSD-A with selections of a new map EBSD-B (Fig. 5). Firstly, the machine learning model is trained with the data from EBSD-A. Afterwards the machine learning model is tested with an entirely new dataset from EBSD-B, which is obtained from a different sector, but from the same MnAl-C sample as in EBSD-A with a size of 180×81 µm² (600×270 pixels). For this new data we compare micromagnetic simulations with the predictions of the machine learning model. We use the same sampling method as from EBSD-A. On a regular grid every second selection with a size of 10×10 pixels is picked and a finite element mesh is created for the simulations. Duplicate selections are removed from the dataset. Figure 5a shows the micromagnetically computed nucleation fields H_n of EBSD-B (about 400 simulations).

Now we can predict the nucleation fields H_n on the very same selections by using the trained regression trees from the original data of EBSD-A. We use the feature vector with the highest precision (see Table 1): GRAIN, SIZE, SW, and TWIN. Figure 5b illustrates those predicted nucleation fields H_n, which show less variation in H_n than the simulated map. Consequently the MAE is larger than in previous results of EBSD-A with a value of 0.157 T/μ_0 (see Table 1). In Fig. 5c we show the absolute deviation $|\Delta H_n|$ of the computed and predicted nucleation fields. There are regions where the machine learning model nicely predicts the computed results, whereas in other regions relevant training data is missing. More data need to be produced to better model these under-represented areas. And probably a large number of grains in the
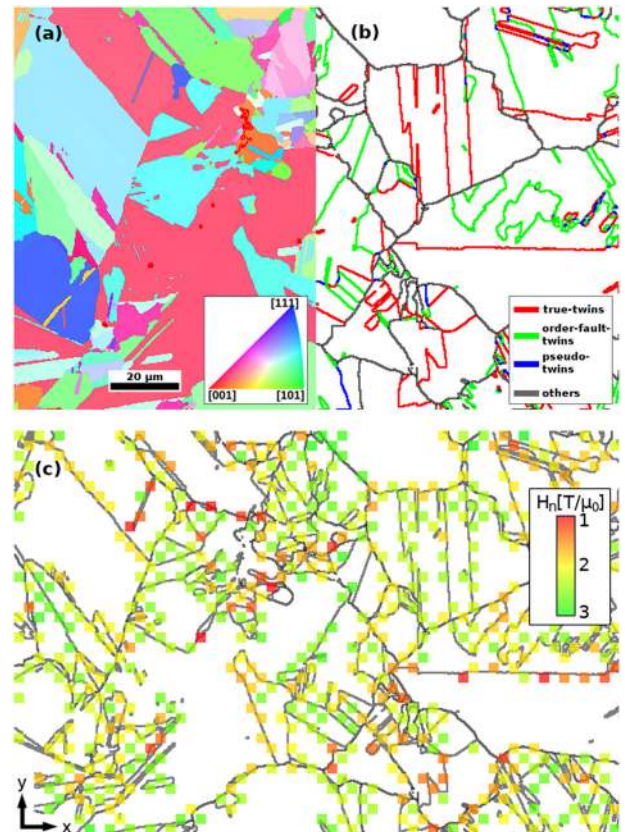
selections causes some of the deviations as well, which is discussed in section Microstructural complexity.

We created a residual plot according to ref. [27], where we show the difference $H_{n,comp} - H_{n,pred}$ versus the predicted nucleation fields $H_{n,pred}$ (Fig. 6). Most residuals lie in a range of ± 1 T/μ_0 and show no trend in the distribution, which is an important indicator for the quality of the machine learning model. In this work we are interested in weak regions of the microstructure, hence absolute values are not that important. With the help of the machine learning model, trends in the distribution of the nucleation fields can be obtained very fast.

Bulk magnetic properties

A trained machine learning model has the ability to predict the local nucleation fields of a full EBSD map within seconds (see EBSD-A in Fig. 2). From the prediction information we suggest a method to construct the hysteresis curve of the whole EBSD sample. In Fig. 7 we show a comparison of two demagnetization curves computed from predictions of EBSD-A with two different feature sets (shown in Fig. 2). We construct the curve with a decreasing external field from 0 to -3 T/μ_0. At each field value we count the selections with a predicted nucleation field greater and lower than this particular field value. As soon as the field value is greater than a locally predicted nucleation field, we assume that
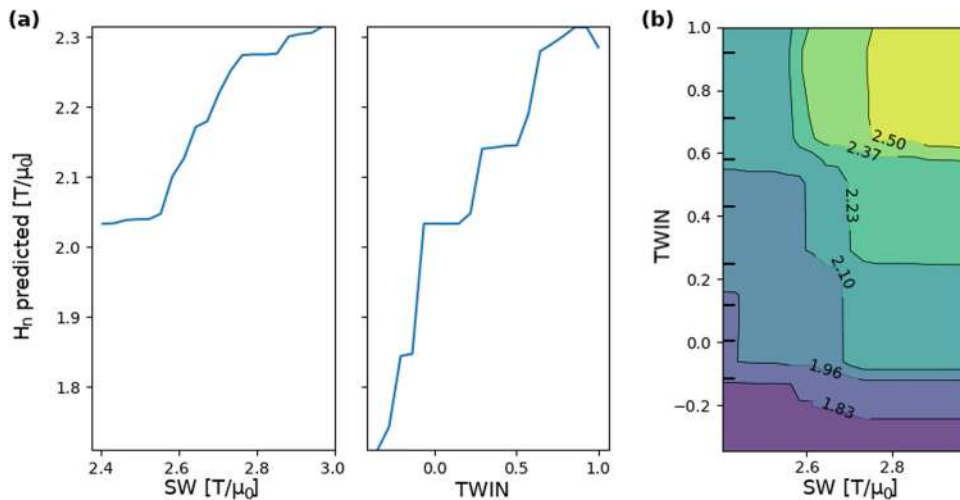
Fig. 4 Partial dependence plots. a One-way partial dependence plots of the features SW and TWIN showing the effect on the predictions. **b** The interaction of SW and TWIN is shown in a two-way partial dependence plot (contour lines with predicted H_n). The model was trained with a feature vector containing all features (PIXEL, GRAIN, SIZE, SW, and TWIN).

Table 2.	Linear regression scores.				
EBSD-A		EBSD-B		Active features	
MAE	**RMSE**	**MAE**	**RMSE**	**SW**	**TW**
0.268	0.348	0.287	0.361	×	
0.214	0.286	0.194	0.280		×
0.186	0.258	0.166	0.245	×	×

Mean absolute error (MAE [T/μ_0]) and root mean squared error (RMSE [T/μ_0]) from the linear regressor are given for the trained models with the features SW and TWIN. The multiplication sign (×) marks the activated feature(s). Results for EBSD-A (Fig. 3) are obtained from the average of a 5-fold cross-validation (Fig. 1a). The results for EBSD-B (Fig. 5) show the performance of the models on the test set after training the model with selections from EBSD-A (Fig. 1b).

this selection is magnetically reversed (switched). With the ratio from unswitched selections to all selections normalized from 1 to −1 we obtain the full curve. Due to the simplification we see a rather high coercive field for both curves ($H_c \approx 2.35\ T/\mu_0$). In reality the coercivity is much lower[6]. In the future we plan to include stray field and exchange coupling between the selections to improve the predicted hysteresis curves.

DISCUSSION
We have demonstrated a fast way to predict local nucleation fields for MnAl-C. An automated script scans a large EBSD dataset, selects hundreds of unique samples and creates quasi-3D finite element meshes accordingly. For each selection the local nucleation field H_n is computed by a fast Landau–Lifshitz–Gilbert t micromagnetic solver. The microstructural features of the EBSD selections and its nucleation fields are used to train random forest and gradient boosting regressors. A voting regressor combines those two methods for an improved prediction accuracy. We use a 5-fold cross-validation to test the performance of single features and feature combinations on an EBSD map. We are using features like PIXEL, that are a set of multiple features, the data points of the EBSD map. On one hand, training the machine with PIXEL includes the orientation of the grains, the geometry as well as the location information. On the other hand, quite a lot of redundant information is in the dataset. The GRAIN
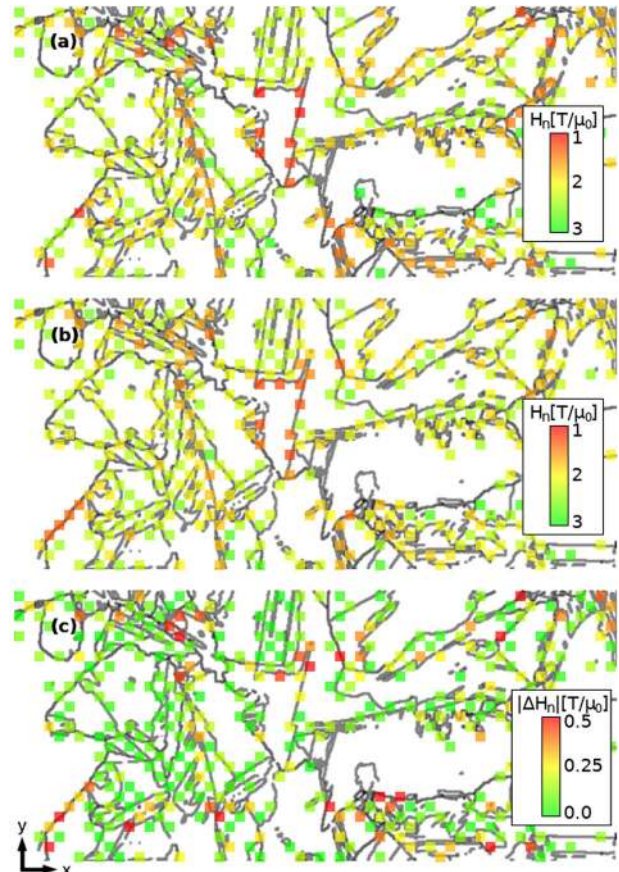

Fig. 5 Comparison of computation and prediction for EBSD-B. EBSD map of 600×270 pixels with 0.3 μm pixel edge length. **a** Nucleation fields H_n computed and **b** predicted by using the data from EBSD-A (Fig. 3) with the features GRAIN, SIZE, SW, and TWIN. **c** Absolute deviation between calculated and predicted nucleation fields.

feature includes only the orientation of the single grains and has no geometrical nor location information. This might be the reason that none of the single or combined features reach a mean absolute error below 0.1 T/μ_0.

**Table 3.** Prediction accuracy according to the number of grains in the selections.

# grains	# selections	MAE [T/μ_0]	RMSE [T/μ_0]
1	46	0.035	0.053
2	316	0.132	0.268
3	150	0.161	0.219
>3	101	0.188	0.246

The scores are obtained by the voting regressor (VR). GRAIN, SIZE, SW, and TWIN are used as training features. The data of EBSD-A (Fig. 3) are split firstly according to the number of grains and than randomly into 80% training data and 20% test data.
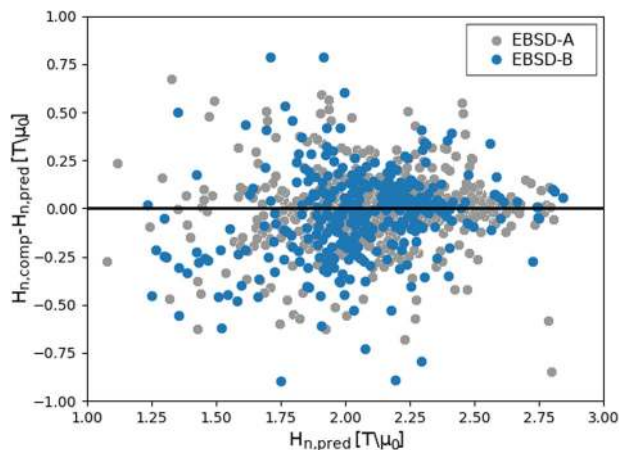
Fig. 6 Residual plot. Residuals are given for predictions of the nucleation fields in EBSD-A (Fig. 3) and EBSD-B (Fig. 5). For the values of EBSD-A we use the validation set of each fold, whereas for EBSD-B we use the full dataset (see Fig. 1a, b, respectively). The model for predicting the nucleation fields used GRAIN, SIZE, SW, and TWIN as features.

The trained machine learning model from the initial map EBSD-A is used to predict local nucleation fields of a new map EBSD-B. Comparison between predictions and computations show that nucleation fields in large areas of the new EBSD map can be nicely reconstructed. Some regions show larger deviations, which might be caused by missing relevant training data and the complexity of the microstructure. Partial dependence plots and a linear regression model showed the strengths of the features SW (minimum Stoner–Wohlfarth switching field) and TWIN (measuring the maximum misalignment between adjacent grains in a selection). Both are derived features based on the physics of magnetization reversal. The feature SW is the classical minimum switching field according to the Stoner–Wohlfarth theory. Exchange interactions between neighboring grains reduce the switching field with increasing misalignment angle[28].

Predicting the nucleation fields on a new map with the linear regressor (active features: SW and TWIN) showed a very good result, coming close to those values, predicted with the best feature combinations of the regression decision trees. Even for differing resolutions of the EBSD datasets the machine learning model could be used, because the geometrical information is not absolutely necessary. In such a case, the error caused by the different pixel size need to be quantified. Using the trained machine learning model a large EBSD dataset can be fully scanned and a map of local nucleation fields can be created within seconds. This information can serve as an input for constructing a hysteresis curve of the whole EBSD sample.
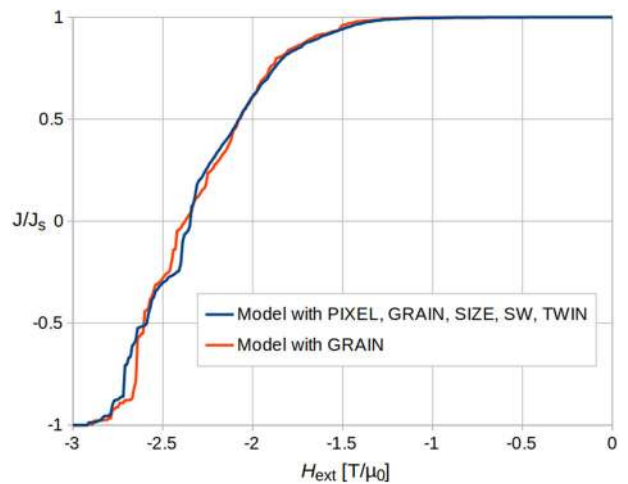
Fig. 7 Full EBSD-A hysteresis curves. Constructed hysteresis curves neglecting stray fields and exchange coupling between selections. The local nucleation fields are taken from Fig. 2a (model with PIXEL, GRAIN, SIZE, SW, and TWIN) and Fig. 2b (model with GRAIN).

Machine learning for magnetic materials faces the following challenges:

(1) Usually machine learning is related to big data whereas in material science the amount of data is typically low. Since it is difficult to get lots of experimental datasets, the use of micromagnetic simulations is a way to generate training data. For the future, we hope that data assimilation methods that combine computed and experimental data will form the basis for datasets that relate coercivity and microstructure. The term data assimilation refers to any systematic procedure that uses experimental measurements to actually improve model simulation[29]. For example, Matsumoto et al. merges ab initio simulations and experimental data to develop light-rare-earth-based permanent magnets[30].

(2) Micromagnetic simulations with the currently available input and computational resources are not accurate enough to give quantitative predictions of nucleation fields and coercivity. Nevertheless, it can be used to see trends in the nucleation field distribution or to find weak spots in the microstructure.

In the following we give prospective applications of machine learning in magnetism:

(1) Machine learning can be used to build better models. It can be seen as a reduced order model that spans over various length scales.

(2) The proposed method can be used to identify weak spots in a magnets microstructure.

(3) Most significant features can give a handle to improve the development process of permanent magnets.

Machine learning in material science is commonly used to characterize the materials' microstructure. In this work we hope to bridge the gap between machine learning and the magnetism community and to foster research in magnetism.

METHODS
In the following section we explain the basics of machine learning and the regression decision trees we used. We give insights in the selected microstructural features and our optimized labels for training the machine learning model. The input data are generated by an automated micromagnetic meshing and simulation routine. It consists of many small selections of a large Electron Backscatter Diffraction (EBSD) image. For each selection we compute the local nucleation field with an external applied

field in y-direction (see coordinate system in all the figures). To improve the prediction accuracy we optimize the hyperparameters of the decision trees.

Machine learning

Machine learning is a statistical approach to automatically analyze large datasets (explained in detail for instance in ref. [24]). The dataset is structured in features and labels. Features are the input for the machine learning model in form of a vector. Labels are the quantities, which we want to teach the model to predict. A model is fitted to the data (training), while it is tested with a subset of the data to get the predictive performance (validation). The trained and validated model is than used to predict the results of new datasets (test). We apply supervised learning where the input data are already given with the true solutions (labels). In this work we use microstructural features of MnAl-C obtained by Electron Backscatter Diffraction (EBSD) microscopy. We generate many small quasi-3D sub-models (selections) of the EBSD dataset and compute the magnetic nucleation fields H_n, the corresponding labels.

In a first step we create input data from EBSD-A (Fig. 3), where we apply a 5-fold cross-validation (Fig. 1a). Training and testing is applied five times with different subdivisions of the dataset, and the prediction accuracy is averaged to obtain an overall training performance. In a second step the machine learning model can be trained again with the full data of EBSD-A (Fig. 1b). For testing we apply this model to EBSD-B (Fig. 5), which is obtained from a different sector, but from the same MnAl-C sample. The data associated with this new EBSD map were not seen before by the machine learning model. For quantitative analysis of the predicted values we are using the mean absolute error (MAE) and the root mean squared error (RMSE). When using the RMSE, large errors are penalized considerably, while using the MAE makes the results easier to interpret.

Classification and regression are the most commonly used tasks in supervised learning. While the labels are split into classes (classification problem), the real values are trained and predicted with regression machine learning models. We use regression decision trees to predict the nucleation field H_n. With linear regression the data are assumed to correspond linearly, yet in real case scenarios, which are usually much more complex, labels depend nonlinearly on the features. Similar to Exl et al.[18], we use random forest (RF) and gradient boosting (GB) decision trees to account for the nonlinearity, which we observed in our results. RF trees combine predictions of individual decision trees trained over randomly generated sub-training samples (often called as ensemble learning). GB trees combine again predictions from several generated sub-training samples, but this time each sample tree learns from residual errors from the previous one. A voting regressor (VR) is then used to average the individual predictions on the whole dataset and to form a final prediction.

We are using Python with the Scikit-Learn framework[24,31] to model the decision trees and to have the respective scores, hyperparameter optimization and visual representations available. The hyperparameters of our RF and GB regressors are optimized with the GridSearchCV method including the full feature vector (see details in section Feature engineering) and a 5-fold cross-validation (see Fig. 1a). As a score we use the negative mean squared error. Details on the optimized hyperparameters can be found in Supplementary Note 1. We tested also a Bayesian optimizer[32] with 5-fold cross-validation to tune the hyperparameters for each possible feature combination (see section Feature importance by feature vector), yet it showed no significant effect on the results. In other words, our regressors are very robust when changing the hyperparameters. Therefore we stick with the hyperparameters from the GridSearchCV approach for the rest of this work.

Dataset generation

Machine learning usually requires a large amount of data to accurately predict new datasets. Using micromagnetic simulations, a reasonable amount of data can be generated for training our regression decision trees. We are computing the nucleation fields H_n of MnAl-C selections obtained from EBSD data. The original data are represented as crystallographic orientations on a regular grid, which are split into two files. One file contains the grain index of each pixel, and in the second the corresponding crystallographic orientation matrices are stored. We translate each orientation matrix into a unit vector, since we calculate micromagnetic simulations on uniaxial grains. EBSD-A (Fig. 3) has a size of $180 \times 120\,\mu m^2$ (600×400 pixels). Since this sample is far too large to simulate, we will simulate about 600 10×10 pixel selections of the entire

image. These selections are picked on a regular grid. In order to limit the number of simulations, only every second selection is taken. We extrude and triangulate each such selection into a finite element mesh with elements much smaller than a pixel and determine their nucleation fields by micromagnetic simulations.

All selections taken from inside the same grain, i.e., the selections include no boundaries and have exactly the same magnetocrystalline orientation, result in the same nucleation field. This is obvious because for all such selections the features are the same. Thus taking more than one selection from the interior of one grain just duplicates the datasets. Here we want to understand how the local structure influences the local nucleation field. Therefore, we can drop the duplicates. Indeed, tests showed the following: Picking randomly 1500 selections from EBSD-A (Fig. 3) without removing duplicates, the prediction accuracy was poor. Sampling on a regular grid with removing duplicates gave the best prediction accuracy. This is consistent with a Monte Carlo study of the influence of duplicate records on regression estimates[33]. Duplicates were found to increase the RMSE depending on their number and position. Figure 3c shows the grain boundaries of EBSD-A and the simulation selections (about 600) colored by their computed nucleation field H_n. Due to the uniqueness constraint, the selections are preferentially located at the grain boundaries. Since nucleation usually starts near defects, which for MnAl-C are most often grain or twin boundaries (Fig. 3b), those locations are the most important determining factors for the bulk coercivity.

Feature engineering

We are using supervised learning, which means, that the input data consists of features and labels (see section Machine Learning). As stated in ref. [34], feature engineering is an essential step in establishing a machine learning model. It is a process of transforming the raw data into good quality features, which improves the overall performance of the machine learning model. Features depend a lot on the underlying problem and on the raw data. In our case the features are extracted from the microstructure of MnAl-C, whereas the labels are the result of micromagnetic simulations (nucleation field H_n).

For each selection of the large EBSD image we know the crystallographic orientation. We can use each individual pixel of an EBSD selection, with its orientation, as a feature and/or we can use each grain, with its orientation. While pixels also include the location coordinates, this information is not given in the list of grains. In our micromagnetic simulations we assume grains with uniaxial magnetocrystalline anisotropy. Hence, it is sufficient to define the magnetocrystalline orientation for each grain as unit vector that points in the hemisphere of the positive y-direction (Fig. 3). The size of the grains can affect the nucleation field and is therefore also considered as a microstructural feature.

The coercivity of permanent magnets is often described by

$$H_c = aH_{N,min} - N_{eff}M_s, \tag{1}$$

where $H_{N,min}$ is the minimum switching field of misoriented grains[8]. The coefficient a expresses the reduction in coercivity due to defects and intergrain exchange interactions. The microstructural parameter N_{eff} is related to the effect of the local demagnetization field near sharp edges and corners of the microstructure. By using the Stoner–Wohlfarth (SW) model for single-domain ferromagnets[35] we analytically calculate the minimum switching field overall grains i:

$$H_{N,min} = \min_i \left\{ 2K/J_s \left(\cos(\beta_i)^{2/3} + \sin(\beta_i)^{2/3} \right)^{-3/2} \right\}, \tag{2}$$

with the misorientation angle β_i between the easy axis of a grain i and the external field. We assume that the external field is applied parallel to the y-axis of a Cartesian coordinate system (Fig. 3). The lowest field value of all grains gives the irreversible switching event for this selection, if other effects like stray field and grain boundary exchange coupling are neglected. The minimum switching field $H_{N,min}$ is used as another feature called SW.

Twin boundaries are known to have a strong influence on the quality of permanent magnets. Especially in MnAl-C true-twins, order-fault-twins and pseudo-twins are frequently found[6], which have a clearly defined misorientation angle. In previous work we computed hysteresis curves for selections of MnAl-C and showed a significant effect of twinning boundaries on the coercive field[14,23]. From manual observation of the simulation results and from partial dependence plots in section Feature importance by partial dependence plots we observe a negative influence

on the coercivity, if the adjacent grains' orientations have a large enclosed angle. Therefore we incorporate the twin angles into a feature for machine learning by calculating the minimum dot product of adjacent grains' orientation unit vectors. A large angle implicitly means that both of the adjacent grains are disadvantageously orientated to the applied external field direction. Whereas the SW model incorporates only the misorientation angle of a single grain's orientation with respect to the external field.

From the microstructural attributes and analytical equations we acquire the following features used for machine learning. In brackets the short feature name is given as well as the number of features as seen by the machine.

- Crystallographic orientation for each data point as unit vector (PIXEL, 300)
- Crystallographic orientation for each grain as unit vector (GRAIN, 60)
- Size of grains (SIZE, 20)
- Minimum Stoner–Wohlfarth switching field $H_{N,min}$ (SW, 1)
- Minimum dot product of adjacent grains' orientation (TWIN, 1)

All the given features form the feature vector **F**:

$$\mathbf{F} = [\underbrace{P_{1xyz} \ldots P_{100xyz}}_{\text{PIXEL}} \underbrace{G_{1xyz} \ldots G_{20xyz}}_{\text{GRAIN}} \underbrace{SI_1 \ldots SI_{20}}_{\text{SIZE}} \text{SW TWIN}] \tag{3}$$

PIXEL's features are three components of the orientation unit vector for each of the 10×10 pixels. The Scikit-Learn software requires a fixed feature vector length, so we set an arbitrary maximum number of grains to 20, which is higher than the maximum number of grains in our selections. Therefore SIZE has 20 entries in the feature vector. And GRAIN has 20 times three entries, which are the components of the orientation unit vectors. We pad the remaining GRAIN and SIZE entries with zeros to fill up the required feature vector length.

Each small selection of the large EBSD image is using the feature vector **F** as input for learning. It is possible to use only a single feature or feature combinations, but the feature vector length must be kept equal within a training set. In other words, our machine learning model is always trained with a single training set with a well defined set of features. Typical feature correlation or feature elimination methods cannot be used in such a configuration, because each component of the unit vectors in PIXEL and GRAIN are already a feature, hence cannot be reasonably compared to other features, like SW or TWIN. In models based on decision trees feature scaling is less important. However, we found that some features contribute more to a change in the nucleation field H_n than others. Feature importance will be discussed in the Results section.

Labeling

Training the machine using supervised learning requires a well defined set of features as well as a label. We defined the microstructural features in the previous section. For the label we chose the nucleation field H_n defined as the value of the external field at which the first irreversible switching event occurs along the demagnetization curve. Irreversible switching can be recognized by a kink in the demagnetization curve, where a discontinuity of magnetization and susceptibility occurs[10]. Often, the switching field is sufficiently defined by the coercivity, i.e., the field which is needed to reverse the sign of the magnetization projected on the field direction as for example the green line in Fig. 8. For such a curve the magnet is clearly in a reversed state after the step. The blue line is much flatter and shows three small steps.

There might be more than one irreversible switching event along the demagnetization curve of a selection. Therefore we have different possibilities in defining a label. It could be either the value of the external field at the (a) first step, (b) the last step, or (c) the largest step in the curve. In order to find the label with the best prediction performance we created a simple classification problem, which tries to predict the 20% percentile of the lowest switching fields. We obtained 90, 85, and 86% correct predictions for using (a), (b), or (c) as label, respectively. Therefore we decided to stick to the first step of the curve (Fig. 8a), the nucleation field H_n, as definition of our label.

Micromagnetic simulations

We generate the input data for our machine learning model with micromagnetic simulations (see section Dataset generation). The resolution of the EBSD dataset is not high enough to be directly used for micromagnetic simulations. Grain boundaries, which are small compared to the grid resolution are not accurately spatially resolved. Sharp angles
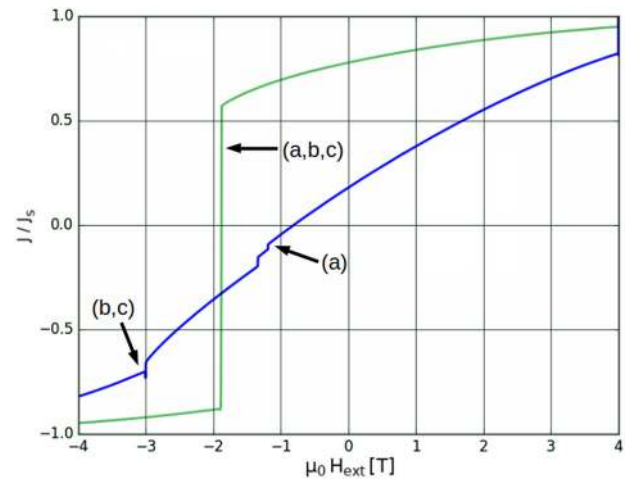

Fig. 8 Sample hysteresis curves. Hysteresis curves may contain single (green) or multiple (blue) switching events. **a** First step, **b** last step, and **c** largest step of the curve.

need to be smoothed for the simulations in order to reduce numerical instabilities. In previous work we showed an automated meshing procedure for the fast generation of simulation models[23]. We combine multiple steps of the toolchain into a single meshing tool using the pre and postprocessing software Salome (salome-platform.org, last visited on 29/07/2019). We load the pixelated EBSD data into the software, smooth the grain boundaries and create a high quality finite element mesh (Supplementary Algorithm 1). The software basically finds all intersection points of the grain boundaries and reconnects them with Bezier curves, replacing the stepped lines of the original pixelated EBSD image. Figure 9 shows an example of an original EBSD selection, the creation of a smoothed 2-dimensional surface and the extruded 3-dimensional mesh. The smoothing operation creates many more data points (pixels) compared to the original selection. A higher number of data points increases the number of components in the feature vector **F** (see section Feature engineering), which slows down the learning process, but does not significantly improve the machine learning accuracy. Therefore we stick to the discretization points of the original EBSD data for training the machine learning model.

The finite element mesh obtained from the automated meshing routine is used by a hybrid finite element boundary method for magnetostatics[17]. We are computing magnetization reversal curves, solving the Landau–Lifshitz–Gilbert (LLG) equation. Intrinsic magnetic properties listed in Table 4 are obtained from bulk MnAl-C[36] at 300 K. The largest edge length of the tetrahedral finite elements, i.e., mesh size, is set to be smaller than the smallest characteristic length of the material[23]. Bance et al. downsized the original EBSD dataset to reduce computational costs by 1:500 and 1:50[14]. Here we use a scaling ratio of 1:15, so that 1 μm in the original EBSD dataset corresponds to 67 nm in the simulations. Please note that scaling might only change the magnetostatic interactions, whereas the influence of misalignment and exchange interactions between different grains on the nucleation is fully taken into account. The final mesh of a single selection has a size of $200 \times 200 \times 40$ nm^3 with about 370,000 tetrahedrons, consisting of 30,000 triangles or 70,000 nodes. A simulation on a single CPU takes several hours depending on the magnetic configuration.

The nucleation fields H_n, obtained by the micromagnetic simulations, are the labels for the machine learning model. We compute the hysteresis curves with an external field aligned in y-direction (Fig. 9) from 4 to −4 T/μ_0 with a sweep rate of 40 mT/ns and a Gilbert damping constant of 1. For each selection we apply a free boundary, which creates demagnetizing fields and may also cause edge nucleation. However we have discovered, that misalignment between neighboring grains (twins) leads to a much larger reduction of coercivity than demagnetizing fields at the edges. The reversed domains nucleate at the interface between strongly misoriented regions (twin boundaries). In other words, the effect of the demagnetizing field is smaller than that of misoriented grains (see section Feature importance by partial dependence plots). Nevertheless we are clear that this gives only an approximation of the local nucleation field. Each squared selection in Fig. 3c is color-coded according to its nucleation field H_n from
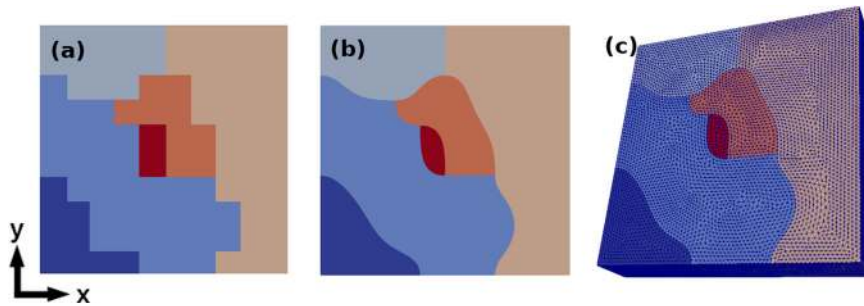
Fig. 9 Automated mesh generation. a Original EBSD data selection, **b** smoothed grain boundaries, and **c** tetrahedral finite element mesh with arbitrary colors according to the grain id. Simulation parameters are given in Table 4.

Table 4.	Micromagnetic parameters for the simulations.	
Parameter	Value	Unit
Temperature T	300	K
Magnetic polarization J_s	0.8	T
Exchange constant A	19.9	pJ/m
Uniaxial anisotropy constant K	1.5	MJ/m^3
External field H_y	4 to -4	T/μ_0
Sweep rate	40	mT/ns
Gilbert damping constant a	1	
Mesh size	3	nm
Scaling ratio	1:15	
Selection size	10×10	Pixels
Extrusion thickness	2	Pixels

Intrinsic magnetic properties J_s, A, and K are obtained from bulk MnAl-C experiments[36].

red (1 T/μ_0) to green (3 T/μ_0). Please be aware of our definition of H_n being the first step in the demagnetization curve as explained in section Labeling.

DATA AVAILABILITY
All data generated or analyzed during this study are available from the corresponding author upon reasonable request.

CODE AVAILABILITY
All the code programmed during this study is available from the corresponding author upon reasonable request.

REFERENCES
1. Skokov, K. & Gutfleisch, O. Heavy rare earth free, free rare earth and rare earth free magnets-vision and reality. *Scripta Materialia* **154**, 289–294 (2018).
2. Coey, J. Permanent magnets: plugging the gap. *Scripta Materialia* **67**, 524–529 (2012).
3. Landuyt, J. V., Tendeloo, G., Broek, J., Donkersloot, H. & Zijlstra, H. Defect structure and magnetic properties of MnAl permanent magnet materials. *IEEE Trans. Magn.* **14**, 679–681 (1978).
4. Houseman, E. & Jakubovics, J. Domain structure and magnetization processes in mnal and mnalc alloys. *J. Magn. Magn. Mater.* **31**, 1005–1006 (1983).
5. Yanar, C., Radmilovic, V., Soffa, W. A. & Wiezorek, J. M. Evolution of microstructure and defect structure in l10-ordered manganese aluminide permanent magnet alloys. *Intermetallics* **9**, 949–954 (2001).
6. Bittner, F., Schultz, L. & Woodcock, T. G. Twin-like defects in L10 ordered τ -MnAl-C studied by EBSD. *Acta Materialia* **101**, 48–54 (2015).
7. Palanisamy, D., Raabe, D. & Gault, B. On the compositional partitioning during phase transformation in a binary ferromagnetic mnal alloy. *Acta Materialia* **174**, 227–236 (2019).
8. Kronmüller, H. & Goll, D. Micromagnetism of advanced hard magnetic materials. *Int. J. Mater. Res.* **100**, 640–651 (2009).
9. Fischbacher, J. et al. Micromagnetics of rare-earth efficient permanent magnets. *J. Phys. D Appl. Phys.* **51**, 193002 (2018).
10. Schabes, M. E. Micromagnetic theory of non-uniform magnetization processes in magnetic recording particles. *J. Magn. Magn. Mater.* **95**, 249–288 (1991).
11. Schwartz, A. J., Kumar, M., Adams, B. L. & Field, D. P. in *Electron Backscatter Diffraction in Materials Science*, vol. 2 (Springer, 2009).
12. Sepehri-Amin, H., Ohkubo, T. & Hono, K. Micromagnetic simulations of magnetization reversals in nd-fe-b based permanent magnets. *Mater. Trans.* **57**, 1221–1229 (2016).
13. Sepehri-Amin, H., Ohkubo, T., Gruber, M., Schrefl, T. & Hono, K. Micromagnetic simulations on the grain size dependence of coercivity in anisotropic nd-fe-b sintered magnets. *Scripta Materialia* **89**, 29–32 (2014).
14. Bance, S., Bittner, F., Woodcock, T. G., Schultz, L. & Schrefl, T. Role of twin and anti-phase defects in MnAl permanent magnets. *Acta Materialia* **131**, 48–56 (2017).
15. Exl, L. et al. Preconditioned nonlinear conjugate gradient method for micromagnetic energy minimization. *Comput. Phys. Commun.* **235**, 179–186 (2019).
16. Tsukahara, H., Iwano, K., Mitsumata, C., Ishikawa, T. & Ono, K. Micromagnetic simulation for the magnetization reversal process of nd-fe-b hot-deformed nanocrystalline permanent magnets. *AIP Adv.* **7**, 056234 (2017).
17. Suess, D. et al. Time resolved micromagnetics using a preconditioned time integration method. *J. Magn. Magn. Mater.* **248**, 298–311 (2002).
18. Exl, L. et al. Magnetic microstructure machine learning analysis. *J. Phys. Mater.* **2**, 014001 (2018).
19. Dimiduk, D. M., Holm, E. A. & Niezgoda, S. R. Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering. *Integ. Mater. Manufact. Innovation* **7**, 157–172 (2018).
20. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. & Walsh, A. Machine learning for molecular and materials science. *Nature* **559**, 547–555 (2018).
21. Azimi, S. M., Britz, D., Engstler, M., Fritz, M. & Mücklich, F. Advanced steel microstructural classification by deep learning methods. *Sci. Rep.* **8**, 2128 (2018).
22. Orme, A. D. et al. Insights into twinning in mg az31: A combined ebsd and machine learning study. *Comput. Mater. Sci.* **124**, 353–363 (2016).
23. Gusenbauer, M. et al. Automated meshing of electron backscatter diffraction data and application to finite element micromagnetics. *J. Magn. Magn. Mater.* **486**, 165256 (2019).
24. Géron, A. in *Hands-on Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* ("O'Reilly Media, Inc.", 2017).
25. Reif, M. & Shafait, F. Efficient feature size reduction via predictive forward selection. *Pattern Recogn.* **47**, 1664–1673 (2014).
26. Hastie, T., Tibshirani, R. & Friedman, J. in *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer Science & Business Media, 2009).
27. Fox, J. in *Applied Regression Analysis and Generalized Linear Models* (Sage Publications, 2015).
28. Schrefl, T., Schmidts, H., Fidler, J. & Kronmüller, H. The role of exchange and dipolar coupling at grain boundaries in hard magnetic materials. *J. Magn. Magn. Mater.* **124**, 251–261 (1993).
29. Abramowitz, G. et al. Neural error regression diagnosis (nerd): a tool for model bias identification and prognostic data assimilation. *J. Hydrometeorol.* **7**, 160–177 (2006).
30. Matsumoto, M. et al. Optimal uni-axial ferromagnetism in (la, ce) _2 fe _{14} b for permanent magnets. *arXivpreprint*: http://arXiv.org/abs/arXiv:1901.10119 (2019).
31. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Machine Learn. Res.* **12**, 2825–2830 (2011).

32. Head, T. et al. scikit-optimize/scikit-optimize: v0. 5.2. github.com/scikit-optimize/scikit-optimize (2018).
33. Sarracino, F. & Mikucka, M. Bias and efficiency loss in regression estimates due to duplicated observations: a monte carlo simulation. *Survey Res. Methods* **11**, 17–44 (2017).
34. Sarkar, D., Bali, R. & Sharma, T. in *A Problem-Solvers Guide To Building Real-World Intelligent Systems* (Apress, Berkely, 2018).
35. Stoner, E. C. & Wohlfarth, E. A. A mechanism of magnetic hysteresis in heterogeneous alloys. *Phil. Trans. R. Soc. Lond. A* **240**, 599–642 (1948).
36. Thielsch, J., Bittner, F. & Woodcock, T. G. Magnetization reversal processes in hot-extruded -MnAl-C. *J. Magn. Magn. Mater.* **426**, 25–31 (2017).

AUTHOR CONTRIBUTIONS

M.G., T.G.W., and T.S. conceived the idea for the present work. M.G. and H.O. carried out all numerical calculations. M.G., H.O., J.F., A.K., and T.S. implemented the automated meshing toolchain, the machine learning framework as well as the micromagnetic solver. P.Z. and T.G.W. acquired and analyzed all experimental data. All authors discussed the results, commented on, and wrote the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at https://doi.org/10.1038/s41524-020-00361-z.

Correspondence and requests for materials should be addressed to M.G.

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.