Materials Discovery 12 (2018) 20–29

Contents lists available at ScienceDirect

Materials Discovery

journal homepage: www.elsevier.com/locate/md

Evolution of MG AZ31 twin activation with strain: A machine learning study

Andrew D. Orme[a,*], David T. Fullwood[a], Michael P. Miles[b], Christophe Giraud-Carrier[c]

[a] Department of Mechanical Engineering, Brigham Young University, 435 CTB, Provo, UT, United States
[b] Department of Manufacturing Engineering, Brigham Young University, 265 CTB, Provo, UT, United States
[c] Department of Computer Science, Brigham Young University, 3328 TMCB, Provo, UT, United States

ARTICLE INFO

Keywords:
Magnesium AZ31
Twin nucleation
Machine learning
Strain dependence
Decision trees
EBSD

ABSTRACT

Complex relationships between microstructure and twin formation in AZ31 magnesium are investigated as a function of increasing strain using supervised machine learning. In one approach, strain is incorporated as an implicit attribute in a single predictive model, in a second method, separate decision trees are formed for each strain level. A comparison of the methods shows that the second better uncovers the underlying physics. The correlations revealed are found to exhibit similarities with parameters used in conventional modeling techniques, leading to the conclusion that machine learning has potential to assist in future microstructural modeling.

1. Introduction

Magnesium (Mg) has played only a minor role in structural and transportation applications, despite its desirable high strength-to-weight ratio. The low ductility of magnesium alloys effectively limits their application in many situations [1–4]. The hexagonal close packed (HCP) structure of magnesium, combined with the strong basal texture that results from sheet rolling processes, limits most readily activated slip to 2 independent basal slip systems at room temperature. In general, 5 slip systems must be active in a material for it to accommodate arbitrary strains [5]. As such, other deformation methods are often activated in the microstructure to fulfill the Taylor criterion, which outlines requirements for accommodation of strain within a polycrystal.

Strain accommodation via tension and compression twinning often provides the necessary deformation modes in HCP materials. While permitting deformation, twin formation can also promote crack formation, leading to material failure [6,7]. Better understanding of microstructure characteristics that influence twin formation in magnesium will assist in controlling this vital deformation mechanism and possibly expand the applications in which magnesium may be used.

Observing and studying twin activity is time consuming. Twin formation often occurs at low frequencies in a microstructure, making it difficult to observe large numbers of twinning instances. Recent advances in electron backscatter diffraction (EBSD) technology and methodology have increased the speed at which microstructure can be observed. When combined with high-resolution EBSD (HR-EBSD, also known as cross-correlation EBSD) methods, the amount of available data on twin activity in Mg alloys is growing rapidly [6–11].

With availability of data increasing steadily, the obstacle becomes efficient and meaningful analysis, as conventional methods for analysis often cannot handle the large quantities of data that emerge from an HR-EBSD scan. These HR-EBSD datasets are ideal for exploration with machine learning and data mining algorithms, which can expose complex and intricate correlations in large datasets with limited prior knowledge (although some prior knowledge can be helpful in tuning the algorithm).

A previous study used this approach to uncover correlations between twin activity and microstructure features in a large dataset of EBSD data collected from samples of Mg AZ31, a common alloy of magnesium [12]. This study by Orme et al investigated twin formation at 2.5% strain, focusing mainly on the effects of microstructure attributes on twin nucleation and propagation.

In the current paper, we apply machine learning to microstructure data taken at several strain steps in an attempt to conduct two investigations. The first is to determine if machine learning can teach us anything new about twin activity in magnesium, with a particular focus on evolution of twin activity with strain. This is accomplished by comparing evolving twin activity characteristics revealed by machine learning with typical saturation-type methods of throttling twin activity in other magnesium modeling approaches. The results from machine learning, which includes strain as an attribute, give insight into twin saturation and provide verification of other parameters used in magnesium modeling.

The second investigation analyzes the machine learning process itself, proposing methods to robustly conduct a machine learning study. Several ways for including strain in the dataset will be explored, with

* Corresponding author.
E-mail address: andreworme@byu.net (A.D. Orme).

https://doi.org/10.1016/j.md.2018.09.002
Received 14 May 2018; Accepted 24 September 2018
Available online 28 September 2018

the merits of each described. The value of machine learning processes for future microstructure data interrogation is also discussed.

1.1. Modeling twinning in magnesium

The complex correlations between microstructure and twinning have been investigated through various means. Statistical methods which involve hypothesis generation and validation have uncovered several key relationships. Work by Barnett et al showed that twinning follows a Hall-Petch type dependence, with large grain sizes yielding higher levels of twinning [13]. Further work by Chino et al supported twinning dependence on grain size by suggesting that twinning depends more on the grain size than on slip orientation [14]. These often form a sort of saturation parameter in modeling, as with increasing strain, large grains available for twinning often are less prevalent, having been divided by previous twin activity or slip.

Twin nucleation is often considered a stochastic event, the local microstructural conditions that determine twin formation are either too complex to incorporate into an efficient deterministic model, or have not all been properly defined yet. Major statistical analysis of twin formation in Mg AZ31 has been performed by Beyerlein et al based upon EBSD data [15,16]. It was found that although twinning does show a dependence on several key microstructure characteristics, such as grain size, grain boundary dislocations, and grain orientation, these attributes do not fully determine twin formation. Rather, they contribute to a statistical probability of twin formation. Twinning was additionally investigated in relation to grain size, grain orientation, grain boundary length, and neighbor misorientation [15].

Work done by Khosravani et al investigated twin formation and nucleation using HR-EBSD data [9]. Twins were found to form at grain boundaries due to slip ('slip assisted') or due to twin propagation over a grain boundary ('twin assisted'). It was shown that slip assisted twins generally formed at high angle boundaries with accumulation of a slip systems in neighboring grains. Twin assisted twins generally formed at low angle boundaries, where grain boundary type was amenable to transmission of a twin from a neighboring grain.

While the statistical studies have been hugely beneficial to improving understanding of twinning in Mg, they generally required the respective researchers to approach the data with pre-determined hypotheses. Modeling was driven by investigation of hypothesized correlations between microstructure characteristics and twinning. An alternative technique to exploring the data is to use machine learning methods. These methods approach the data in a different way, where a large number of attributes suspected to have correlations with twinning can be investigated at the same time. The machine learning algorithms sort through the data and return only relevant correlations, resulting in knowledge discovery. Correlations not previously considered by the researcher can be readily revealed by the algorithms [17].

1.2. Machine learning for data investigation

With rapidly growing datasets on material structure and behavior, data mining and machine learning are starting to be used more frequently in materials science applications [12,17–28]. Sometimes referred to as "materials informatics," this relatively new technique for materials characterization and analysis has proven very beneficial. Bostanabad et al used machine learning techniques for the creation of statistically similar microstructures for use in modeling applications [20]. Others have used machine learning to classify, identify and design appropriate microstructures [18,21,22,26]. Machine learning is helping the materials science community to accelerate knowledge discovery and knowledge application [26,18–28].

Various types of machine learning have been employed in these studies. Supervised machine learning methods rely on a set of attributes with a defined class, or outcome. These learning algorithms form classification- or regression-type correlations between the attributes and the class. Unsupervised machine learning methods do not have a class attribute supplied to them, rather, they use clustering techniques to determine correlations among attributes. This study focuses on applying supervised machine learning algorithms, specifically a decision tree classifier. This classifier produces an easily interpreted decision structure, and is outlined in greater detail in the methods section.

Preliminary work in this area applied machine learning techniques to investigate correlations between twin activity and microstructural features, at a single level of compressive strain [12]. The study validated correlations previously uncovered via statistical methods (such as the influence of grain size and grain boundary type on twin formation) and expanded them to include other attributes (such as basal Schmid factors and local dislocation density). However, twin formation is clearly not a static or isolated event, as the deformation level increases various factors disrupt or evolve the interrelations between microstructure and twin activity. These include twin saturation, texture evolution, secondary twin formation, strain hardening behavior and various other mechanisms. In this study, machine learning techniques are applied to investigate how the roles of various microstructural factors that influence twin behavior are modified with strain level, from a purely data analysis point of view (i.e., from an observation rather than a predictive view). Strain is included in the machine learning models in two methods, and the merits of each are evaluated.

2. Materials and methods

2.1. Sample preparation

The material used for this study is a commercially available 3mm thick cold rolled and annealed AZ31 Mg alloy sheet. These sheets exhibit a strong basal texture, which is ideal for twin formation. Before testing, the sheet was determined to be twin free using optical microscopy. Test specimens were cut from the sheet at a size of 3mm x 4mm x 3mm using a diamond blade saw. Specimens were then mechanically polished using abrasive sheets and diamond abrasive pastes. Etching to prepare the samples for study using an electron microscope was performed using a solution of 60% ethanol, 20% water, 15% acetic acid, and 5% nitric acid. A focused ion beam was used to place fiducial marks on the surface of the specimens to allow for approximate 2D strain measurements in the scan area [9].

A single specimen was strained to levels of ~2.5%, ~5%, and ~7.5% via compression of the sample along the RD. While micro strain values may differ, the macro strain is assumed to be fairly accurate. This assumption is valid as long as deformation across the sample is homogeneous; if bulging occurs at the sample edges then the true strain value may be different. Multiple areas of the sample surface were scanned at each strain level to assure representative data was collected for the microstructure. EBSD data for the microstructure was collected using an S-FEG XL30 FEI SEM with an EDAX EBSD camera connected to OIM Data Acquisition software. Areas of approximately 90 μm by 140 μm were scanned at a step size of 300nm. Sample areas at each strain level are included in Fig. 1. Following data collection, scan data was processed using OIM Analysis, MATLAB, and Open XY (HR-EBSD) to prepare attributes for machine learning [8,29].

For analysis with machine learning, one attribute had to be identified as the class attribute, or the effect of interest to the study. The remaining attributes constitute the possible causes of the effect. Twin formation is the phenomenon of interest in this study and was selected as the class attribute. Twin formation is assumed to have occurred in any grain containing a twin. Characteristics of grains containing twins are used to attempt to define parameters influencing twinning. Twins identified for this study are either tensile twins (of the $\{10\bar{1}2\}\langle 10\bar{1}\bar{1}\rangle$ variants) or compression twins (of the $\{10\bar{1}1\}\langle 10\bar{1}\bar{2}\rangle$ variants). These were identified using automated methods included in OIM Analysis [30]. An example of twin identification by OIM is included in Fig. 2. The twins observed in this study were predominantly of the

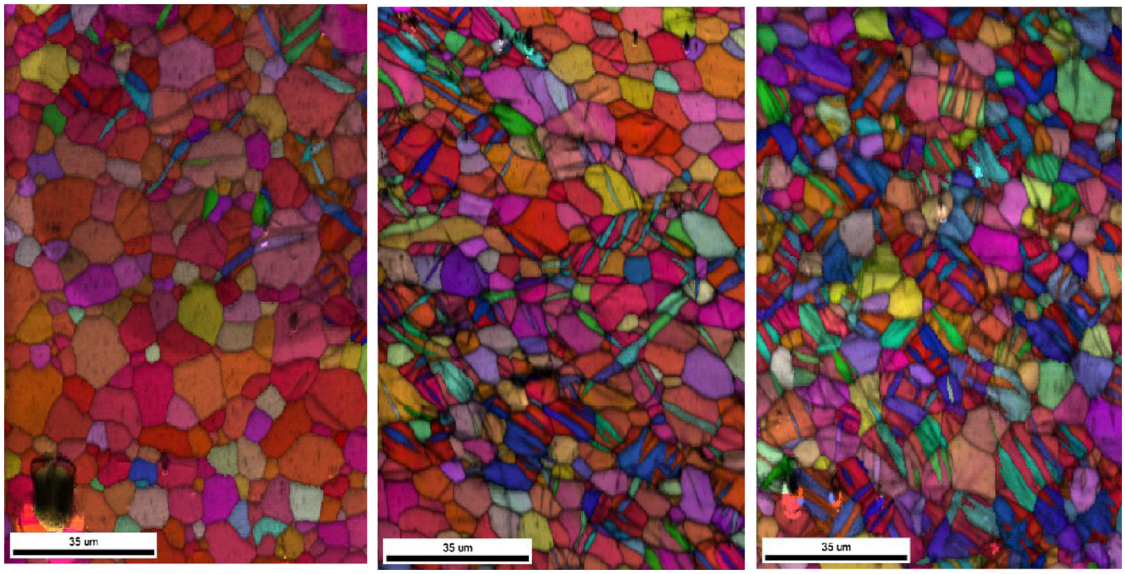

Fig. 1. Inverse pole figure/confidence index maps of scans. From left to right strains are 2.5%, 5%, 7%. Increasing twinned area should be noted for greater strain level.

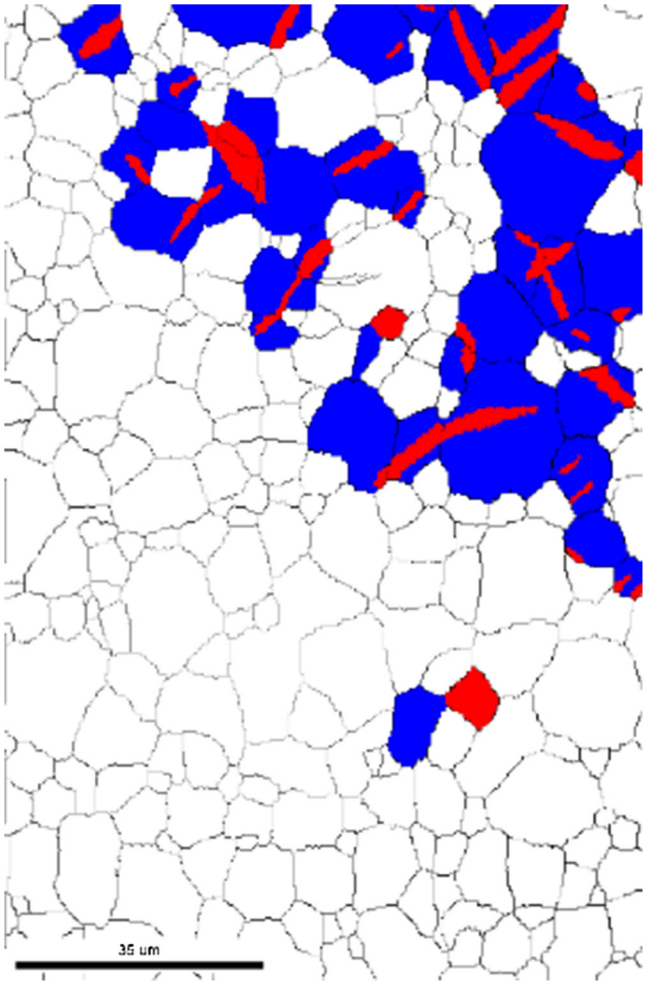

Fig. 2. Twin identification example. The red regions are identified by OIM as twins (both compression and tension varieties). The blue regions are the parent grains (grains in which there is twin activity) and are the grains labeled as having formed twins in the dataset used for machine learning. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

$\{10\bar{1}2\}\langle 10\bar{1}\bar{1}\rangle$, or tensile, variants, and thus only attributes connected to tensile twinning were considered. Attributes considered for potential correlation with twin activity are described in Table 1. The complete dataset consisted of data for approximately 7000 instances, where each instance is a grain.

Two methods for including strain information into the machine learning model were used. The first involved the creation of one large dataset (including all 7000 instances) which contained strain as an attribute. In this method, strain appears in the resulting model decision structure. The location of strain-based decisions in the tree relative to other attributes provides insight into its importance and highlights connections between strain, other attributes, and twinning.

The second method broke the large dataset into 3 smaller datasets (of ~2300 instances each), containing instances from only one strain level (2.5%, 5%, or 7%). These models do not include strain as an attribute. Correlations between twinning and strain are gained by comparing differences and similarities between the three trees, noting shifts in tree structure and attribute thresholds (for example, highlighting whether there is some critical grain size for twin formation that changes with increased strain). Decision trees of both types were created using the methods outlined in the following section and were compared to evaluate which method most efficiently provides information about twinning dependence on strain.

2.2. Machine learning

Following the EBSD data collection and preprocessing to extract attribute values, datasets were loaded into the WEKA machine learning workbench, distributed for public use by the University of Waikato [31]. Machine learning algorithms in WEKA were used to uncover relationships between the attributes in Table 1 and twinning events. Of special interest was the possible evolution of these relationships with increasing levels of strain. The following methodology for machine learning was used for our research and is presented in detail for reference in future materials science machine learning applications. It follows a traditional machine learning process, wherein data is preprocessed, then analyzed using inductive learning algorithms, and finally the induced models are evaluated.

2.2.1. Preprocessing

Data pre-processing algorithms, also known as filters in WEKA, were used to prepare datasets for input into machine learning algorithms.

Table 1
Attributes included in project datasets. Note the attribute set is similar to that used by Orme et al. [12] but now includes strain as an attribute.

Attribute	Abbreviation	Description
Strain	STRAIN	Nominal compression strain
Grain size	SIZE	Equivalent diameter. Calculated as the diameter of a circle with the same area as the measured grain
Neighboring grain size	NBRSIZE	Average neighbor grain size
Relative grain size	RELSIZE	Grain size divided by neighboring grain size
Number of neighbors	NUMNBRS	Number of neighboring grains
Deviation of c-axis from RD, TD, and ND	RDMISO, NDMISO, TDMISO	Smallest angle of misorientation
Kernel average misorientation	MISOGRAIN	Average misorientation of directly neighboring points (with 5° cutoff)
Schmid factors (SF) of basal $< a >$ and pyramidal $< c+a >$ slip systems	BASALSF, CASF	Maximum value of each slip system taken as a grain average,
Schmid factors of $\{10\bar{1}2\}\langle 10\bar{1}\bar{1}\rangle$ tensile twins	TWINSF	Maximum value, taken as a grain average (also considers the possibility of negative values)
Ratio of twinning Schmid factor to $< c+a >$ Schmid factor	TWINCASF	Maximum SF for twinning divided by max SF for pyramidal $< c+a >$
Local dislocation densities	LOGDD	Grain average of sum of Nye tensor terms
Twin in grain (1 if true, 0 if false)	TWINID	Class attribute indicating presence of a twin in a grain

Since our target attribute consists of nominal, rather than numeric, values, our problem is a classification problem. A study where the target attribute is numeric is generally handled using a regression approach. Input attributes that take non-continuous values, including discrete numbers (e.g., number of neighbor grains), were transformed into nominal values. Machine learning algorithms handle continuous and discrete attributes differently, hence attributes must be correctly designated as either nominal or continuous.

Overfitting is a significant risk in machine learning. An overfit model is one that has been trained in such a way that it has a developed bias towards the data used to train. This is often identified when a model induced by a machine learning algorithm exhibits high accuracy on its training data, but performs significantly worse on previously unseen test data. One way to detect overfitting is to use a "data hold out" method. We remove and save a portion of the training data to be used as a validation set. The remaining data is used for training and inducing a model. Once formed, the model is then tested against the validation set. If the accuracies obtained on the training data and on the validation data are similar, the model is deemed to not be overfit. Here, we select 10% of the training data at random for validation. We repeat the process three times with different random samples from the dataset. We only report the results for the first sample since results were comparable across the three hold out samples.

The dataset contained ~5000 instances where grains did not twin with only ~2000 instances where grains did twin. While this is representative of natural twinning occurrences at these strain levels, this skew in the data yielded models that were heavily biased, with low accuracy in predicting twin formation events. In order to reduce the skew of the dataset, a balancing filter was used to weight instances. This filter maintains the distribution of instances for each attribute while equalizing the classes. This is accomplished by randomly removing no-twin instances and adding weights to twin instances. Classes were balanced to approximately 50/50 of twin / no-twin cases, resulting in a model able to predict both outcomes with comparable accuracy.

2.2.2. Feature selection

Feature selection methods were employed to reduce the number of attributes used in training. Feature selection assists the machine learning process by selecting attributes relevant to the desired classification, improving data quality, increasing the speed of the algorithm, removing redundant attributes, and enhancing the interpretability of results [32]. A wrapper subset evaluator was employed to create a subset of attributes to be used for training of the classifier. This method evaluates the merit of several attribute subsets using a learning scheme. Attribute subsets that demonstrate high accuracy with the defined learning scheme are suggested to the user. Here, we select the J48 decision tree classifier as the learning scheme.

For the work presented here, a receiver operating characteristic (ROC) curve evaluation was used as the measure for determining the optimal subset. ROC curves plot the true positive rate on a vertical axis against the true negative rate on a horizontal axis [33]. The area under these curves (AUC) is frequently used to compare classifier performance. A perfectly optimized classifier will have an AUC of 1. The attribute selection method explored the effects of inducing models with different attribute subsets on the ROC curve and returned the attribute set corresponding to the highest AUC value. The algorithm repeated the process 10 times, with varied random seeding each time. The percentage of times each attribute was selected was reported. Attributes selected over 50% of the time were retained in the dataset. All other attributes were removed from the dataset.

2.2.3. Classifier training and evaluation

The subset of attributes was then used to train a J48 decision tree classification algorithm. This algorithm results in easily interpretable decision structures, which take the form of decision trees. A simple decision tree is shown in Fig. 3. This example tree predicts whether it is appropriate to play outside or not based on the current weather. The structure represents a set of consecutive decisions and predicted outcomes based on those decisions. In the tree in Fig. 3 for instance, a decision is made by first considering the outlook, then based on the outlook, considering either the wind or humidity. A sample decision sequence is as follows; if the outlook is sunny and the humidity is normal, it would be appropriate to play outside. As seen, rule statements are easily read off a decision tree. Tree terminology is used to

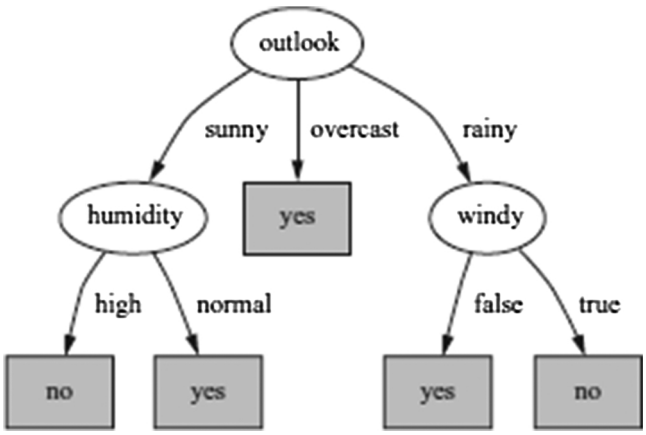

Fig. 3. Sample decision tree. Note how correlations between the weather attributes and the decision are easily read off the tree, e.g., if the outlook is sunny, but the humidity is high, you shouldn't play outside [33].

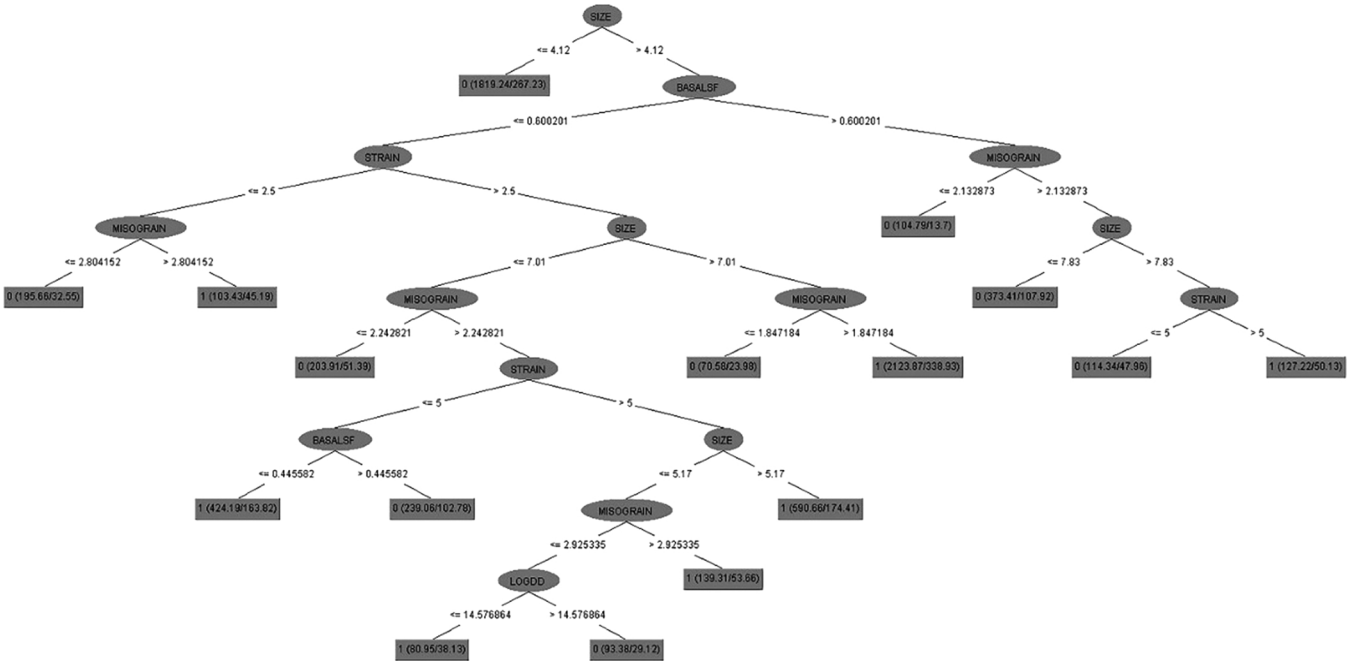

Fig. 4. Combined strain model.

describe the tree. The decision at the top of the tree is called the "root," divisions are referred to as "branches," and final classifications are referred to as "leaves." The J48 algorithm selects the decisions by analyzing a metric called information gain. Information gain is a measure of entropy reduction caused by dividing data. The algorithm evaluates all potential decisions and picks the one with the largest information gain. Divisions appearing higher in the tree inherently have larger information gain and thus can be considered to have the greatest importance in the phenomenon being studied.

10-fold cross validation was employed to assess the accuracy of the models. Similar to the "data hold out" method mentioned earlier, 10-fold cross-validation randomly splits the dataset into 10 equal folds. One at a time, a fold is held out, while the other 9 are used to train the classifier. The held-out fold is tested with the classifier and the accuracy recorded. The process continues with each fold being removed and tested. The final estimate of accuracy is the average over the 10 folds.

Other parameters of importance for the models include the confidence cutoff value and leaf size. These metrics affect the size of the final tree, which directly influence the interpretability of a model. The confidence cutoff controls the pruning decisions made by the classifier. During tree construction, some small decision branches may form that have little impact on the overall tree accuracy. These small branches are removed from the tree during a process called error-based pruning. The amount of pruning is controlled using a confidence value, with lower values yielding higher amounts of pruning. For this study, a moderate level of pruning was desired and a confidence value of 0.25 was selected to accomplish this. The reader is referred to reference [34] for more information on pruning. Pruning trades accuracy for simplicity, which is desired in this application [35]. The leaf size parameter dictates when the algorithm should stop attempting to form a decision split down a branch. It specifies the minimum number of classified instances that need to be present after a decision for a leaf to form. This project used a minimum leaf size of 80 instances, meaning a leaf must contain at least 80 instances to be included in the decision tree. Both the confidence cutoff and leaf size values were determined empirically by comparing the differences in model accuracy with changes in both values.

2.3. No free lunch theorem

Relative to the selection of a given machine learning technique, Wolpert proved what are referred to as the "no free lunch" theorems [36]. These state that a given learning scheme cannot beat random guessing for all possible learning situations. In simpler terms, no one machine learning algorithm will always perform better than others. There is no universally applicable machine learning technique. The selection of a machine learning method is highly dependent on the problem to be investigated.

It should be thus noted, that although the methods outlined above were successful in this study, they may not perform well in all applications. The methods are provided in detail to illustrate principles which govern machine learning decisions. Adaptation or replacement may be required for other applications [33,37,38].

3. Results

3.1. Feature selection results

As noted in section 2, feature selection was performed on the initial set of attributes included in

Table 1. This reduced the complexity of the dataset significantly by reducing the number of attributes, ideally making it easier for the classifier to give clear and accurate results. These attributes were selected by the feature selection algorithm as being the most important independent attributes for twin formation. The final set of attributes determined to be most relevant to twin formation are:

- Strain: nominal compression strain applied to the sample
- Size: equivalent diameter of grain
- Kernel Average Misorientation: average misorientation of neighboring points within a grain
- Basal SF: maximum value of each slip system taken as a grain average
- LogDD: grain average of sum of Nye tensor terms

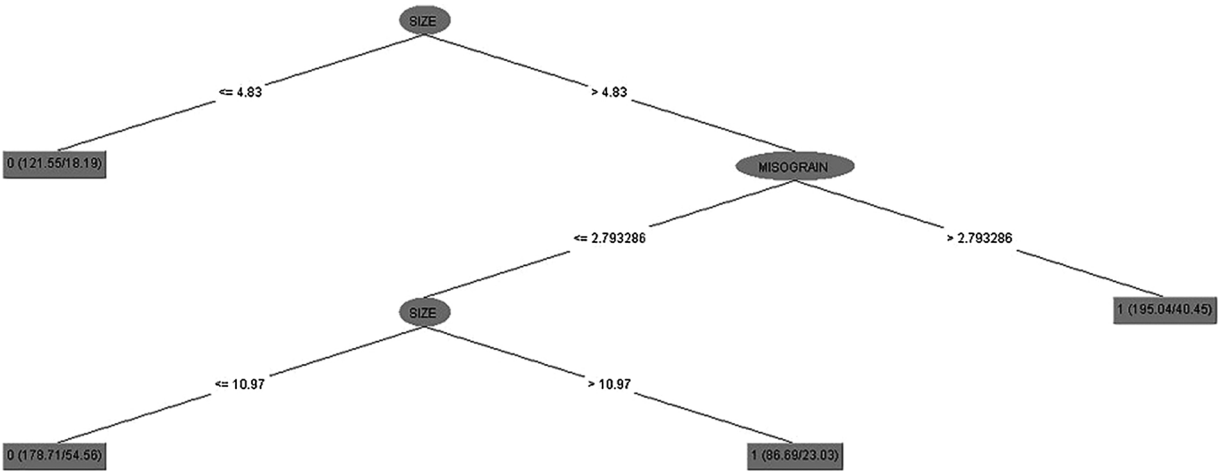

Fig. 5. 2.5% strain model.

3.2. Modeling results

Following the methods outlined above, several decision trees were obtained from the J48 algorithm. Each are included in this section of the paper. Fig. 4 is the model created using the large dataset where strain was an attribute. Figs. 5–7 were created from the smaller datasets, with each tree representing a single strain level. Significant differences in the resulting decision structure for each method were noted and will be explored further. For each tree, decisions rules connecting twinning and strain were extracted. It should be noted that due to the nature of machine learning algorithms, these rules result in probability of twinning and not in certainty of twinning events.

Results provided for each tree include an image of the actual decision structure, an accuracy using cross-validation, an accuracy using holdout data, and an AUC (area under the ROC curve) value. Due to the class balancing perfomed during pre-processing, the reported accuracies should be compared to a default accuracy of 50%, i.e. random classification, calculated as the accuracy when selecting the majority class. The holdout accuracy is a validation that the model has not been overfit to the training data. The AUC value represents the fit of the model to the classification space, with a perfectly fit model having an AUC value of 1. Each tree contains accuracy data for each leaf, or end point in the structure. Each leaf contains a fraction value, the numerator is the weight of instances correctly classified in that leaf, while the denominator is the weight of instances incorrectly classified in that leaf.

3.3. Combined strain model

The tree obtained for training on the dataset containing the selected attributes (including strain) is shown in Fig. 4. This tree had 75.8% accuracy using cross validation and 77.5% accuracy using holdout data. The AUC value is 0.836.

Key decision rules were manually extracted from the tree following model creation.

- If strain is > 5% & basal SF is > 0.6 & kernel average misorientation is > 2.13° & the grain is > 7 μm: twinning will likely occur.
- If strain is > 5% & basal SF is < 0.6 & size is < 7.01 μm: twinning likely will occur if the grain is either > 5.17 μm in size OR has a kernel average misorientation > 2.92° OR has a dislocation density < 14.6.

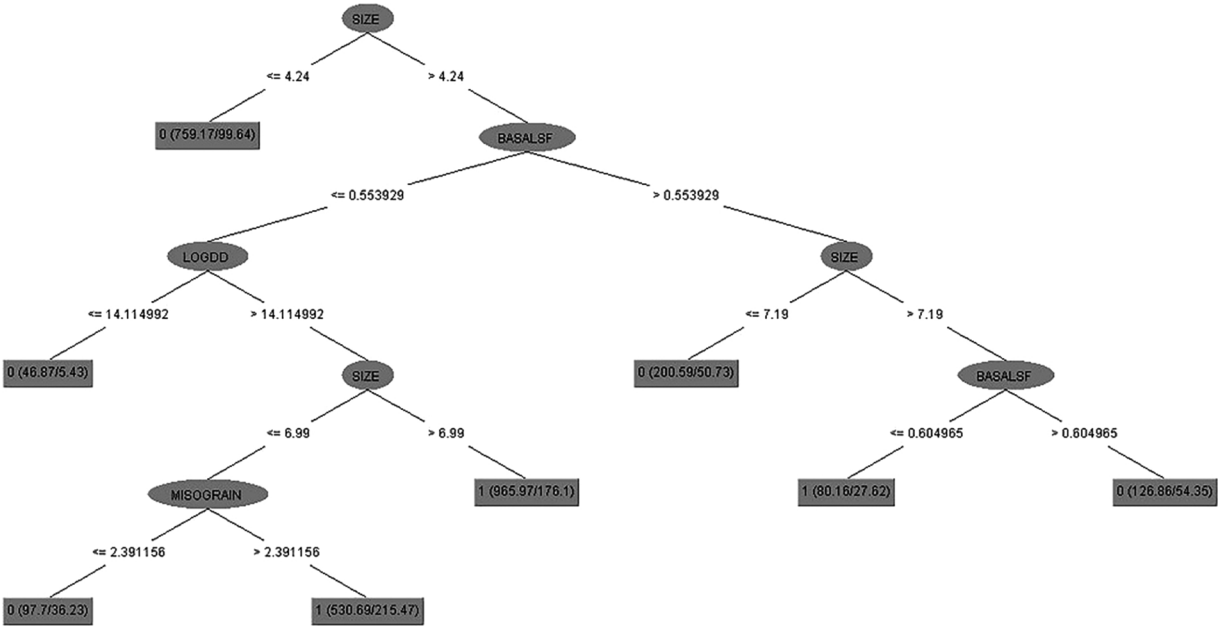

Fig. 6. 5% strain model.

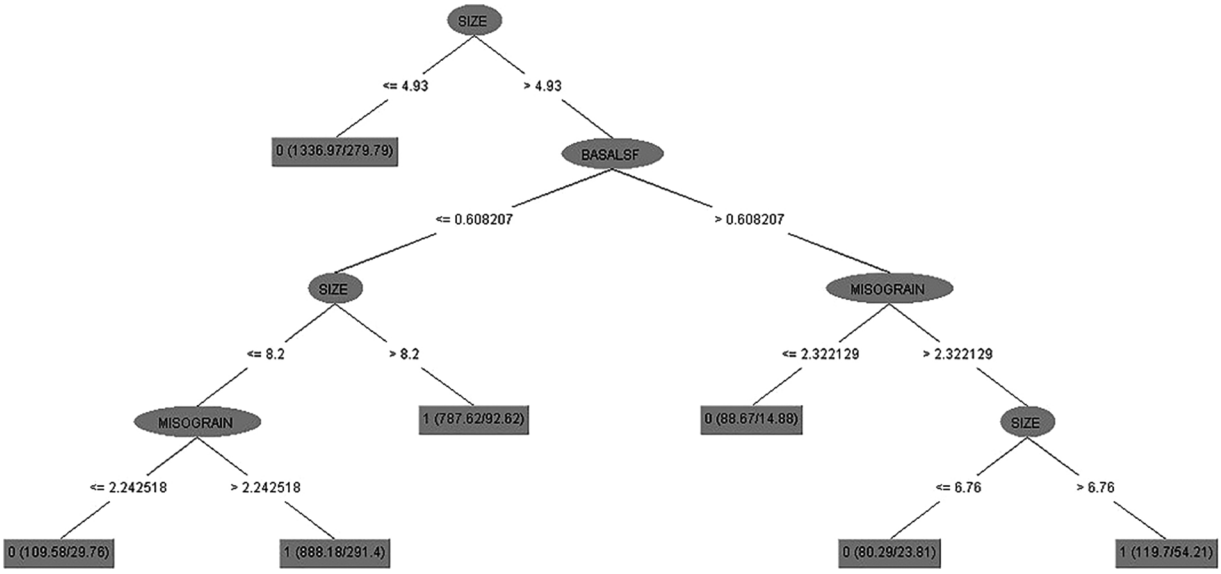

Fig. 7. 7% strain model.

- If basal SF is < 0.6 & strain is > 2.5% & the grain is > 7.01 μm in size & has a kernel average misorientation > 1.85°: twinning will likely occur.
- If strain is between 2.5% and 5% & the grain is < 7.01 μm in size & has a kernel average misorientation > 2.24° & basal SF is < 0.45: twinning will likely occur.
- If strain is below 2.5% & the kernel average misorientation is > 2.8° & basal SF < 0.6: twinning will likely occur.
- In any of these cases, grain size must be greater than 4 μm for a probability of twin formation.

3.4. 2.5% strain model

The tree obtained for training on the dataset containing the selected attributes using only data recorded at 2.5% strain is shown in Fig. 5. This tree had 72.8% accuracy using cross validation and 83.1% accuracy using holdout data. The AUC value is 0.857.

Key decision rules were manually extracted from the tree following model creation.

- If the size is less than 4.83 μm: twinning will not likely occur.
- If the grain size is larger than 4.83 μm & the misorientation is greater than 2.79°: twinning will occur.
- For a misorientation less than 2.79°, the grain must be larger than 10.97 μm for twinning to likely occur.

3.5. 5% strain model

The tree obtained for training on the dataset containing the selected attributes using only data recorded at 5% strain is shown in Fig. 6. This tree had 73.6% accuracy using cross validation and 75.7% accuracy using holdout data. The AUC value is 0.809.

Key decision rules were manually extracted from the tree following model creation.

- Grain must be larger than 4.24 μm for any possibility of twinning activity.
- The grain must be larger than 7.19 μm & have a basal SF between .55 and .6 for possibility of twinning.
- If the Basal SF is less than .55, it must have a dislocation density > 14.1 & large size (> 7 μm) OR large kernel average misorientation (> 2.39°) for possibility of twinning.

3.6. 7% strain model

The tree obtained for training on the dataset containing the selected attributes using only data recorded at 7% strain is shown in Fig. 7. This tree had 75.3% accuracy using cross validation and 77.1% accuracy using holdout data. The AUC value is 0.801.

Key decision rules were manually extracted from the tree following model creation.

- Grain must be > 4.93 μm for the possibility of twinning.
- For possibility of twinning, the size must be greater than 6.76 μm & have a basal SF > 0.6 & a kernel average misorientation > 2.32°.
- If the basal SF is less than 0.6, twinning will likely occur either if the grain size is > 8.2 μm OR if the kernel average misorientation is > 2.24.

4. Discussion

We first explore insights into the machine learning process itself. On average, the models returned approximately 75% accuracy using both cross validation with the training data and with the holdout data. Since the two accuracies (cross validation and hold out) calculated for each model were generally comparable, the “data hold out” method and other methodology outlined in the methods section of the paper appears to have been successful in creating robust models. These accuracies could potentially be increased by expanding the initial dataset to include new attributes or by further refining of the pre-processing methodology.

There are obvious differences in the resulting models created with the two methods for including strain. Including strain as an attribute in the dataset resulted in a complex decision tree (more branches and resulting leaves), as seen in Fig. 4. The trees created by separating the data into distinct datasets based on strain are much simpler (fewer branches and leaves). Table 2 contains the accuracies and AUC for each model. It should be noted that for each performance metric, there is minimal difference between the combined strain model and the strain separated models. Using performance metrics, the two methods used to account for strain in the modeling process are essentially equal in merit. A more qualitative exploration of the trees is required to determine if one method holds greater merit than the other.

The goal of the machine learning modeling completed in this work is to explore variations in attributes with changes in strain. As such, we

Table 2
Performance measures for each machine learning model.

Model	Cross Validation Accuracy	"Hold Out" Accuracy	AUC Value
Combined Strain	75.8%	77.5%	0.836
2.5% Strain	72.8%	83.1%	0.857
5% Strain	73.6%	75.7%	0.809
7% Strain	75.3%	77.1%	0.801

desire models which expose clear correlations between strain and grain attributes. The combined model is a more complex decision structure, making it difficult to recognize clear correlations between strain and attributes. The strain separated models provide a much clearer picture, without a compromise in accuracy.

A key example of this are trends in grain size. The strain separated models exhibit a decrease in grain size associated with twinning at small grain misorientations and basal Schmid factors, from 10.97 μm at 2.5% strain to 8.2 μm at 7% strain. While twinning at these grain sizes is also dependent on other attributes, the change in required grain size is readily noted. This strain dependent trend is not easily revealed from the relationships extracted from the combined strain model, since strain appears several levels into the decision tree, making extraction of such correlations difficult. The tree does not contain comparable relationships at each level of strain. Instead, the combined strain tree offers better insight into activity between levels of strain, for instance between 2.5% and 5% strain, giving thresholds and attribute values that describe twin activity in this region. Since grain size is a parameter of interest in our comparison of saturation methods, the types of clear relationships revealed by the strain separated models provide greater insight for comparisons. This leads us to conclude that, at least for this study, separating the dataset based on strain is the better option.

We now explore whether the models yield insight into the twinning phenomenon in magnesium. Comparisons between the results of this paper and the methods used by others in modeling will be conducted to determine if any new information was uncovered by the machine learning models. To the best of the authors' knowledge, little research has been done investigating shifting relationships between attributes and twin formation over varying strain levels. Extensive work has been done, however, investigating changes in relationships with changing strain rates and temperatures. Additionally, many plasticity models include parameters that have parallels to the attributes used in this study, such as saturation measures. These studies and models will be used to make comparisons with the correlations found in this paper.

The correlations extracted from the trees exhibited several notable trends with increases in strain. Basal Schmid factor threshold values did not demonstrate much (if any) shift with strain. In general, basal SF threshold values stayed around 0.6. Work exploring critical twin stresses and relationships with microstructural attributes shows some promising correlations with these results [39]. As seen in our model, there appears to be a value for basal Schmid factor of the grain which defines whether twinning may or may not occur. This could be referred to as a threshold dividing grains that are "soft," meaning they prevent twin growth in favor for slip, or "hard," meaning they allow twin formation before slip. Hutchinson et al performed a study in which they investigated critically resolved shear stress for twin systems in magnesium, which supports the idea of a threshold between a hard and soft grain [39]. Additional work by Wang et al investigated texture effects that interact with Hall-Petch relationships, mentioned earlier [40].

Grain size thresholds showed significant shift over strain. At low strains, the grains must be approximately 2 μm bigger than at high strains. A universal minimum cutoff for twinning of ~ 4 μm was observed meaning, in general, grains smaller than 4 μm will not twin. This behavior was explored by Meyers et al in their investigation of twin formation [41]. A large dependence between critical twinning stress and grain size was observed in their work. This was expanded to a Hall-Petch relationship, with a kT value defined for hardening due to twins, in addition to a kS value defined for hardening due to slip. As found through the machine learning algorithms, grain size plays a large role in determining twin activity, with differing requirements at varying strain levels [13,42].

In our model, both kernel average misorientation and dislocation density give insights into dislocation activity within the microstructure. Neither exhibited a strong influence on the model, tending towards smaller dislocation values associated with twinning at higher strains.

The paper by Meyers also investigated connections between strain and twinning, most importantly, whether twinning shows a similar level of sensitivity to strain rate and temperature that slip does [41]. It was suggested that, unlike slip, twinning does not have a well-established sensitivity to such external parameters. Twinning may instead exhibit a stronger dependence on internal parameters, like grain size and dislocation density [43]. The proposed lack of twinning sensitivity to external parameters may be one reason why no major shifts in attribute thresholds were noted in this study. Small shifts were observed, but over a rather large strain range, no major changes occurred.

One explanation for differences in twin activity at various strain levels is what Meyers refers to as unloading of stress in twinned regions, resulting in a higher required stress to produce subsequent twin activity [41]. Additionally, at a higher stress, large grains have a high likelihood of already being twinned, meaning available grain area available for twinning is generally smaller, resulting in a sort of strain hardening phenomenon [44]. Smaller grains may have different thresholds required for twinning to occur. Additionally, twinning may appear to exhibit dependencies on external factors, such as strain rate, when in reality the dependence is due to slip [45]. The close and complex relationship between slip and twinning may cause complications in fully understanding unique twinning dependencies.

Of particular interest is how the results obtained in this study compare with typical crystal plasticity models, which are used heavily to model twinning and de-twinning activity in materials such as magnesium [46–48]. These models base prediction of twin activity mainly on critically resolved shear stress (CRSS) for twinning and on a variety of hardening laws which govern twin growth, shrinkage, and saturation. These models also take into account local twin strain, and do so in a wide variety of ways [49]. Several models utilize the Taylor model which assumes plastic strain on a grain is equal to the macroscopically applied strain [50]. Others use self-consistent models where strain varies in each grain and is governed by a relative stiffness as compared to the rest of the material [51–53]. A more advanced model created by Beyerlein et al takes into account additional factors such as dislocation considerations (movement, hardening, temperature dependence) and twin volume fraction [54].

Since most plasticity models include some consideration of twin CRSS, it is interesting to note that the models created in the current investigation did not select twin SF as a contributing attribute. Instead, the models defined twinning activity based on the SF of the basal slip system, which is the other common deformation system in HCP metals [55]. The relationships show that twinning will occur if the SF for a given grain is less than a designated threshold, in essence suggesting that the grain is not well oriented for basal slip. This could be due to a higher occurrence of basal slip at the strains and temperatures observed, causing the basal SF to have higher predictive ability for the data used than twin SF.

In addition, due to the strong basal texture caused by cold rolling, we have in essence constrained the texture of the system to a constant. A representation of this texture is shown in Fig. 8. A result of this strong texture is a low information gain for any attributes containing texture-based information, since differences in texture values (such as SF and CRSS) are small across the scanned areas. To further investigate texture based effects on twinning, data from a randomly textured microstructure would need to be used with the machine learning algorithms.

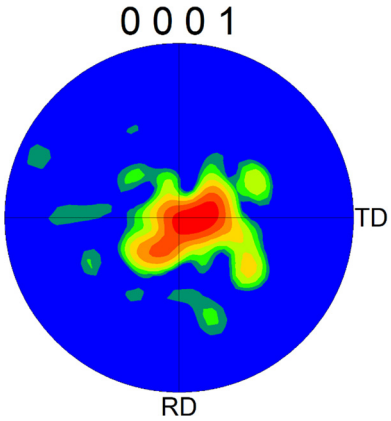

Fig. 8. Starting texture pole figure. Note the strong basal texture exhibited by the magnesium sample.

5. Conclusion

Large datasets containing information about twin formation in MgAZ31 over several strain levels were analyzed using machine learning. Results extracted from decision trees revealed insights into the dependence of attributes on strain level. It was found that for the attributes considered, only grain size requirements showed a heavy dependence on strain. Other factors found to have the greatest impact on twin formation, based on the datasets used in this study, were measures of basal Schmid factor and dislocation density. The thresholds noted for these attributes were compared and found to be supported by previous plasticity modeling of Mg AZ31. It was found that the predictive ability of this model may be limited due to the homogenous texture constraint imposed by using cold rolled specimens, which may account for the lack of strong texture correlations, as compared with other plasticity models. It was suggested that simpler models resulting from separating the dataset based on strain are more useful in gaining insights into twinning than a larger, more complex model. Machine learning was shown to be an effective tool in investigating large materials datasets. Knowledge gained from such studies has the potential to assist in refining twin formation and plasticity models, possibly contributing to an increase in their accuracy.

Declarations of interest

None.

Acknowledgments

The authors would like to acknowledge Ali Khosravani, who oversaw the collection of the raw EBSD data used in this paper. The data he collected has been useful in various projects, including the predecessor to this paper [12].

We would also like to recognize the Brigham Young University Honors Program for their assistance in facilitating this research opportunity and providing guidance and feedback during the writing process.

This material is based upon work supported by the National Science Foundation REU supplement to grant CMMI 1404771. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

[1] T. Al-Samman, G. Gottstein, Room temperature formability of a magnesium AZ31 alloy: examining the role of texture on the deformation mechanisms, Mater. Sci. Eng. A 488 (2008) 406–414.
[2] A. Chapuis, J. Driver, A fundamental study of the high temperature deformation mechanisms of magnesium, J. Phys. Conf. Ser. 240 (2010) 012092.
[3] K. Piao, K. Chung, M.G. Lee, R.H. Wagoner, Twinning-slip transitions in Mg AZ31B, Metall Mater. Trans. A 43A (9) (2012) 3300–3313.
[4] B.L. Mordike, T. Ebert, Magnesium: properties — applications — potential, Mater. Sci. Eng. A 302 (1) (2001) 37–45.
[5] Rv. Mises, Mechanik der plastischen Formänderung von Kristallen, ZAMM-J. Appl. Math. Mech./Zeitschrift für Angewandte Mathematik und Mechanik 8 (3) (1928) 161–185.
[6] M.R. Barnett, Twinning and the ductility of magnesium alloys Part II. "Contraction" twins, Mater. Sci. Eng. a-Struct. 464 (1–2) (2007) 8–16.
[7] M.R. Barnett, Twinning and the ductility of magnesium alloys: part I:' Tension' twins, Mater. Sci. Eng. A 464 (1–2) (2007) 1–7.
[8] BYU, OpenXY, github.com, 2015.
[9] A. Khosravani, D. Fullwood, J. Scott, M. Miles, R. Mishra, Nucleation and propagation of 1012 twins in AZ31 magnesium alloy, Acta Mater. 100 (2015) 202–214.
[10] A. Khosravani, Application of High Resolution Electron Backscatter Diffraction (HR-EBSD) Techniques to Twinning Deformation Mechanism in AZ31 Magnesium Alloy, (2012).
[11] H. Yang, S. Yin, C. Huang, Z. Zhang, S. Wu, S. Li, Y. Liu, EBSD study on deformation twinning in AZ31 magnesium alloy during quasi-in-situ compression, Adv. Eng. Mater. 10 (10) (2008) 955–960.
[12] A.D. Orme, I. Chelladurai, T.M. Rampton, D.T. Fullwood, A. Khosravani, M.P. Miles, R.K. Mishra, Insights into twinning in Mg AZ31: a combined EBSD and machine learning study, Comp. Mater. Sci. 124 (2016) 353–363.
[13] M.R. Barnett, Z. Keshavarz, A.G. Beer, D. Atwell, Influence of grain size on the compressive deformation of wrought Mg–3Al–1Zn, Acta Mater. 52 (17) (2004) 5093–5103.
[14] Y. Chino, K. Kimura, M. Hakamada, M. Mabuchi, Mechanical anisotropy due to twinning in an extruded AZ31 Mg alloy, Mater. Sci. Eng. A 485 (1) (2008) 311–317.
[15] I.J. Beyerlein, R.J. McCabe, C.N. Tome, Effect of microstructure on the nucleation of deformation twins in polycrystalline high-purity magnesium: a multi-scale modeling study, J. Mech. Phys. Solids 59 (2011) 988–1003.
[16] I.J. Beyerlein, C.N. Tome, A probabilistic twin nucleation model for HCP polycrystalline metals, Proc. R. Soc. A 466 (2010) 2517–2544.
[17] A. Jain, G. Hautier, S.P. Ong, K. Persson, New opportunities for materials informatics: resources and data mining techniques for uncovering hidden relationships, J. Mater. Res. 31 (08) (2016) 977–994.
[18] N. Altinkok, R. Koker, Modelling of the prediction of tensile and density properties in particle reinforced metal matrix composites by using neural networks, Mater. Des. 27 (8) (2006) 625–631.
[19] P. Antony, P. Manujesh, N. Jnanesh, Data mining and machine learning approaches on engineering materials—A review, Recent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE International Conference on, IEEE, (2016), pp. 69–73.
[20] R. Bostanabad, A.T. Bui, W. Xie, D.W. Apley, W. Chen, Stochastic microstructure characterization and reconstruction via supervised learning, Acta Mater. 103 (2016) 89–102.
[21] A. Chowdhury, E. Kautz, B. Yener, D. Lewis, Image driven machine learning methods for microstructure recognition, Comp. Mater. Sci. 123 (2016) 176–187.
[22] B.L. DeCost, E.A. Holm, A computer vision approach for automated analysis and classification of microstructural image data, Comp. Mater. Sci. 110 (2015) 126–133.
[23] T. Lookman, F.J. Alexander, K. Rajan, Information Science for Materials Discovery and Design, Springer Series in Materials Science, Springer, Switzerland, 2016, p. 307 225.
[24] N. Lubbers, T. Lookman, K. Barros, Inferring Low-dimensional Microstructure Representations Using Convolutional Neural Networks, arXiv Preprint arXiv:1611.02764, (2016).
[25] T. Mueller, A.G. Kusne, R. Ramprasad, Machine learning in materials science: recent progress and emerging applications, Rev. Comput. Chem. 29 (2016) 186.
[26] G. Pilania, C.C. Wang, X. Jiang, S. Rajasekaran, R. Ramprasad, Accelerating materials property predictions using machine learning, Sci. Rep.-Uk 3 (2013).
[27] K. Rajan, Informatics for Materials Science and Engineering: Data-driven Discovery for Accelerated Experimentation and Application, Butterworth-Heinemann, 2013.
[28] D. Xue, P.V. Balachandran, J. Hogden, J. Theiler, D. Xue, T. Lookman, Accelerated search for materials with targeted properties by adaptive design, Nat. Commun. 7 (2016).
[29] I. The Mathworks, Matlab, The Mathworks, Inc., 2006.
[30] B. Henrie, T. Mason, B. Hansen, A semiautomated electron backscatter diffraction technique for extracting reliable twin statistics, Metall. Mater. Trans. A 35 (12) (2004) 3745–3751.
[31] I.H. Witten, E. Frank, M.A. Hall, C.J. Pal, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2016.
[32] H. Liu, H. Motoda, Feature Selection for Knowledge Discovery and Data Mining, Kluwer Academic Publishers, Massachusetts, 1998.
[33] I.H. Witten, E. Frank, M.A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, Elsevier, 2011.
[34] Bethopedia: Pruning. http://wiki.bethanycrane.com/pruning, 2012 (Accessed May 16 2017.2017).
[35] M. Bohanec, I. Bratko, Trading accuracy for simplicity in decision trees, Mach. Learn. 15 (3) (1994) 223–250.
[36] D.H. Wolpert, The lack of A priori distinctions between learning algorithms, Neural Comput. 8 (7) (1996) 1341–1390.
[37] P. Domingos, A few useful things to know about machine learning, Commun. ACM 55 (10) (2012) 78–87.

[38] S. Chakrabarti, E. Cox, E. Frank, R.H. Güting, J. Han, X. Jiang, M. Kamber, S.S. Lightstone, T.P. Nadeau, R.E. Neapolitan, Data Mining: Know It All, Morgan Kaufmann, 2008.
[39] W. Hutchinson, M. Barnett, Effective values of critical resolved shear stress for slip in polycrystalline magnesium and other hcp metals, Scripta Mater. 63 (7) (2010) 737–740.
[40] Y. Wang, H. Choo, Influence of texture on Hall–Petch relationships in an Mg alloy, Acta Mater. 81 (2014) 83–97.
[41] M.A. Meyers, O. Vohringer, V.A. Lubarda, The onset of twinning in metals: a constitutive description, Acta Mater. 49 (19) (2001) 4025–4039.
[42] A. Jain, O. Duygulu, D. Brown, C. Tomé, S. Agnew, Grain size effects on the tensile properties and deformation mechanisms of a magnesium alloy, AZ31B, sheet, Mater. Sci. Eng. A 486 (1) (2008) 545–555.
[43] M. Lentz, A. Behringer, C. Fahrenson, I.J. Beyerlein, W. Reimers, Grain size effects on primary, secondary, and tertiary twin development in Mg-4 wt pct Li (-1 wt pct Al) alloys, Metall. Mater. Trans. A 45 (11) (2014) 4737–4741.
[44] X. Lou, M. Li, R. Boger, S. Agnew, R. Wagoner, Hardening evolution of AZ31B Mg sheet, Int. J. Plast. 23 (1) (2007) 44–86.
[45] M. Knezevic, M. Zecevic, I.J. Beyerlein, J.F. Bingert, R.J. McCabe, Strain rate and temperature effects on the selection of primary and secondary slip and twinning systems in HCP Zr, Acta Mater. 88 (2015) 55–73.
[46] H. Wang, P. Wu, J. Wang, C. Tomé, A crystal plasticity model for hexagonal close packed (HCP) crystals including twinning and de-twinning mechanisms, Int. J. Plast. 49 (2013) 36–52.
[47] A. Staroselsky, L. Anand, A constitutive model for hcp materials deforming by slip and twinning: application to magnesium alloy AZ31B, Int. J. Plast. 19 (10) (2003) 1843–1864.
[48] S. Agnew, C. Tomé, D. Brown, T. Holden, S. Vogel, Study of slip mechanisms in a magnesium alloy by neutron diffraction and modeling, Scripta Mater. 48 (8) (2003) 1003–1008.
[49] H. Wang, B. Raeisinia, P. Wu, S. Agnew, C. Tomé, Evaluation of self-consistent polycrystal plasticity models for magnesium alloy AZ31B sheet, Int. J. Solids Struct. 47 (21) (2010) 2905–2917.
[50] G.I. Taylor, Plastic strain in metals, Our. Inst. Metals 62 (1938) 307–324.
[51] E. Kröner, Berechnung der elastischen Konstanten des Vielkristalls aus den Konstanten des Einkristalls, Zeitschrift für Physik A Hadrons Nuclei 151 (4) (1958) 504–518.
[52] R. Hill, Continuum micro-mechanics of elastoplastic polycrystals, J. Mech. Phys. Solids 13 (2) (1965) 89–101.
[53] J. Hutchinson, Bounds and self-consistent estimates for creep of polycrystalline materials, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, (1976), pp. 101–127. The Royal Society.
[54] I. Beyerlein, C. Tomé, A dislocation-based constitutive law for pure Zr including temperature effects, Int. J. Plast. 24 (5) (2008) 867–895.
[55] S. Agnew, O. Duygulu, Plastic anisotropy and the role of non-basal slip in magnesium alloy AZ31b, Int. J. Plast. 21 (2005) 1161–1193.