2015-03-01

Deformation Twin Nucleation and Growth Characterization in Magnesium Alloys Using Novel EBSD Pattern Analysis and Machine Learning Tools

Travis Michael Rampton
Brigham Young University - Provo

Deformation Twin Nucleation and Growth Characterization in Magnesium Alloys Using Novel EBSD Pattern Analysis and Machine Learning Tools

Travis Michael Rampton

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

David T. Fullwood, Chair
Brian D. Jensen
Michael P. Miles
Tracy W. Nelson
Eric R. Homer

Department of Mechanical Engineering

Brigham Young University

March 2015

ABSTRACT

Deformation Twin Nucleation and Growth Characterization in Magnesium Alloys Using Novel EBSD Pattern Analysis and Machine Learning Tools

Travis Michael Rampton
Department of Mechanical Engineering, BYU
Doctor of Philosophy

Deformation twinning in Magnesium alloys both facilitates slip and forms sites for failure. Currently, basic studies of twinning in Mg are facilitated by electron backscatter diffraction (EBSD) which is able to extract a myriad of information relating to crystalline microstructures. Although much information is available via EBSD, various problems relating to deformation twinning have not been solved. This dissertation provides new insights into deformation twinning in Mg alloys, with particular focus on AZ31. These insights were gained through the development of new EBSD and related machine learning tools that extract more information beyond what is currently accessed.

The first tool relating to characterization of deformed and twinned materials focuses on surface topography crack detection. The intensity map across EBSD images contains vital information that can be used to detect evolution of surface roughness and crack formation, which typically occurs at twin boundaries. The method of topography recovery resulted in reconstruction errors as low as 2% over a 500 μm length. The method was then applied to a 3 μm x 3 μm area of twinned Tantalum which experienced topographic alterations. The topography of Ta correlated with other measured changes in the microstructure. Additionally, EBSD images were used to identify the presence of cracks in Nickel microstructures. Several cracks were identified on the Ni specimen, demonstrating that cracks as thin as 34 nm could be measured.

A further EBSD based tool developed for this study was used to identify thin compression twins in Mg; these are often missed in a traditional EBSD scan due to their size relative to the electron probe. This tool takes advantage of crystallographic relationships that exist between parent and twinned grains; common planes that exist in both grains lead to bands of consistent intensity as a scan crosses a twin. Hence, twin boundaries in a microstructure can be recognized, even when they are associated with thin twins. Proof of concept was performed on known twins in Inconel 600, Tantalum, and Magnesium AZ31. This method was then used to search for undetected twins in a Mg AZ31 structure, revealing nearly double the number of twins compared with those initially measured by standard procedures.

To uncover the driving forces behind deformation twinning in Mg, a machine learning framework was developed to leverage all of the data available from EBSD and use that to create a physics based models of twin nucleation and growth. The resultant models for nucleation and growth were measured to be up to 86.5% and 96.1% accurate respectively. Each model revealed a unique combination of crystallographic attributes that affected twinning in the AZ31.

Keywords: EBSD, Magnesium, deformation twinning, surface topography, machine learning

ACKNOWLEDGEMENTS

First of all, I would like to acknowledge and thank my graduate committee Dr. David Fullwood, Dr. Brian Jensen, Dr. Michael Miles, Dr. Tracy Nelson, and Dr. Eric Homer for taking the to time review this dissertation.

I would also like to acknowledge financial support from National Science Foundation grant CMMI 1404771 and Department of Energy grant DE-SC0012587 in addition to Dr. Raja Mishra from General Motors for funding this project and providing materials for study.

Thank you to all the members of the lab with whom I had the pleasure of working: Stuart Rogers, Dr. Jay Basinger, Ali Khosravani, Samikshya Subedi, Tim Ruggles, Thomas Hardin, Ribeka Takahashi, Caroline Sorenson, Craig Daniels, and many more.

Dr. David Fullwood deserves an award for his perseverance in seeing that I finish my degree even while working a full-time job through the last 18+ months. I have enjoyed my time with him; he has made me consider things which I couldn't do on my own, and he has truly been an amazing advisor. He saw potential in me a long time ago that I hope to repay with future work.

I want to thank my family for their support, but most of all I want to thank my dear wife for sharing the difficult load of graduate school with me over the last few years. This was a great team effort.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

The dissertation presented here is divided into three sections (chapters 2 - 4) that are written as stand-alone papers for publication purposes. Chapters 2 and 3 develop tools to extract information from electron backscatter diffraction (EBSD) images beyond what is currently measured. This additional data includes surface topography, crack detection and improved compression twin identification. Chapter 4 of this dissertation combines EBSD derived data, including high-resolution EBSD (HR-EBSD) information, from Mg AZ31 samples to identify mechanisms that relate to twin nucleation and growth. Because of the complexity of deformation twinning in Mg AZ31, machine learning is applied to the Mg data to extract basic physics behind twin nucleation and growth. Machine learning is a computer-based technique that excels at finding correlations in complex data, which can be used to create predictive models. In the case of Mg, simplified physical models will be created to explain deformation-twin related activity.

In summary, the second chapter of this paper will apply basic knowledge of electron beam interaction with a surface to reconstruct surface features from unprocessed EBSD images. In the third chapter, the mechanics of EBSD pattern formation will be considered in order to improve the spatial resolution of twins in a microstructure such as Mg AZ31. Finally, the fourth chapter of this dissertation will cover the development of machine learning towards the extraction of twinning mechanisms in Mg AZ31.

The basis of this dissertation relies on three main principles: deformation twinning in Mg, EBSD, and machine learning. As such, an introduction to these topics is provided.

1.1 Deformation Twinning in Magnesium

Before the mechanics of deformation twinning in Mg can be uncovered it is first important to understand the complexity of slip systems required to accommodate plastic deformation in this HCP material (Agnew and Duygulu 2005, Koike 2005, Graff, Brocks et al. 2007, Izadbakhsh, Inal et al. 2011). In Mg, there are several potential slip systems: basal <a> (2 independent), prismatic <a> (2 independent), pyramidal <a> type I (4 independent), and pyramidal <c+a> (5 independent). Furthermore, plastic deformation can be accommodated by compression twinning (mainly the six $\{10\bar{1}1\}\langle 10\bar{1}2\rangle$ variants), and tensile twinning (mainly the six $\{10\bar{1}2\}\langle 10\bar{1}\bar{1}\rangle$ variants). The Taylor model of plastic deformation requires that at least 5 independent slip systems be activated to accommodate an arbitrarily imposed strain (Taylor 1938). While there is a sufficient number of slip systems in Mg to accommodate the Taylor model (minimum of 5), only the basal system and prismatic <a> systems are easily activated, providing 4 independent, active slip systems at room temperature (see Table 1-1). This leaves a requirement of one additional slip system for compliance with the Taylor criterion.

The potential activity of a slip system most often follows a CRSS model. In the case of rolled AZ31 sheet the strong basal texture requires some slip or twinning along the <c+a> direction in order to accommodate contraction or extension of the c-axis (Jonas, Mu et al. 2011).The much higher τ_{crss} values of the <c+a> systems (see Table 1-1 for ranges of values reported in the literature) indicates at a simplistic level, for example, that tensile twinning will occur before <c+a> slip to provide the 5th active system in the Taylor model (B.C. Wonziewiz

1967, Agnew, Yoo et al. 2001, Agnew, Tome et al. 2003, Koike, Kobayashi et al. 2003, Agnew and Duygulu 2005, Koike 2005). While prior observations have provided useful information about the relative difficulties of slip or twinning in various systems, a stochastic approach to the prediction of twin initiation is needed in order to improve models of material behavior for complex strain path deformation, as occurs during industrial forming operations.

Table 1-1: Ranges of published τ_{crss} values for various deformation mechanisms in Mg at room temperature.

Slip System	basal <a>	prismatic <a>	pyramidal <c+a>	$\{10\bar{1}2\}$ twinning	$\{10\bar{1}1\}$ twinning
τ_{crss} (MPa)	4	8-10	80-100	11-12	76-153

As such, there is a need to fill the current void in the detailed connections between meso-scale cause and effect of twinning (Barnett, Keshavarz et al. 2004, Beyerlein, McCabe et al. 2011, Izadbakhsh, Inal et al. 2011). Work by Barnett et al showed that twin nucleation in Mg follows a Hall-Petch relationship, where the required twinning stress increases with smaller grain size (Barnett, Keshavarz et al. 2004). Furthermore, Beyerlein et al recently combined atomistic simulations with an extensive EBSD-based study to demonstrate the effects of grain boundary (GB) misorientation and GB dislocation structure on the nucleation of twins (Beyerlein, McCabe et al. 2011). The resultant model proposed a stochastic approach to twin nucleation, combined with a CRSS basis for twin propagation. It incorporated distinct weights for the probability of twin nucleation on grain boundaries above and below 45^{o}misorientation, due to the observed tendency of twins to be present at low angle GBs. A similar study by Khosravani (Khosravani, Fullwood et al. Submitted 2012) further categorized twin nucleation events at GBs into spontaneous formation of twins (slip-assisted) and propagation of twins across grain boundaries (twin-assisted). The importance of dislocation structure near GBs was also highlighted. The

study further demonstrated that twins easily propagate through low angle (15-25^{o}) GBs and tend to nucleate at high angle GBs (>39^{o}). The different considerations of each of these models might be reconciled into one model by taking a different approach in which large data sets are explored using machine learning to reveal correlations as the basis for model structure and parameters.

1.2 Electron Backscatter Diffraction (EBSD)

Crystallographic information, such as twinning, can be measured in Mg using modern EBSD techniques. EBSD provides the ability to extract a myriad of metrics for input into a machine-learning framework. The automated acquisition of EBSD data has been used in materials science for several decades, during which time the process has achieved common data collection speeds of hundreds of points per second. Such speeds allow for relatively large microstructures to be quickly and accurately measured (0.3^{o} resolution in lattice orientation). The gathered EBSD data can then be processed by commercially available software to produce other crystallographically significant data related to grain orientation, phase, and morphology (OIM). Available information from traditional EBSD techniques also includes grain size distribution, Schmid factors, variations (gradients) in lattice orientation, and GB misorientations.

Additionally, in recent years high resolution EBSD (HR-EBSD) techniques have been developed to extract even more information from the collected EBSD data (Wilkinson, Meaden et al. 2005, Villert, Maurice et al. 2009, Gardner, Adams et al. 2010, Basinger, Fullwood et al. 2011). These methods apply cross-correlation to EBSD images in order to measure orientations and relative rotations with even greater angular resolution (0.006^{o}). Using the cross-correlation technique, HR-EBSD is also capable of being used to measure (relative) elastic strain and geometrically necessary dislocation (GND) density at each data point. Both measurements rely on the accurate extraction of the deformation gradient tensors as well as curvatures. Curvature

measurements in particular are applied toward the calculation of the Nye dislocation tensor which can then be used to estimate the GND content at a given sample point. These additional tools help provide dislocation data over large sample areas that are available through SEM methods. While HR-EBSD measurements of dislocation density are still in the development stages, they have been shown to provide reliable results (Wilkinson, Meaden et al. 2006, Landon, Adams et al. 2008, Kacher, Adams et al. 2009, Britton and Wilkinson 2011).

With the combination of HR-EBSD data and standard EBSD metrics the mechanisms involved in deformation twinning of Mg can be more fully characterized. The incorporation of standard EBSD data will include among other metrics grain size, local orientation gradients, orientation of the c-axis relative to the sheet normal direction, grain boundary misorientations, and the Schmid factors for <a> type slip, <c+a> type slip, and twinning. From the author's point of view these metrics among a few others constitute a broad set of crystallographic measurements that may relate to the nucleation and propagation of twinning in AZ31.

1.3 Machine Learning

With the abundance of crystallographic information that may affect twinning in AZ31 and given the uncertainty underlying the cause and effect of twin nucleation and propagation, a modeling approach that minimizes the assumptions made about the nature of the events under investigation would be useful for gaining insights into the actual causes of these events. Machine learning, including various types of data mining developed to find statistical correlations among large datasets, offers one such method of non-discriminatory characterization (Yoram and Nahum , Reich and Travitzky 1995, Sha and Edwards 2007, Tompos, Margitfalvi et al. 2007). The basic idea of machine learning is to create relationships between input parameters and observed output by analyzing individual events or occurrences and generalizing similarities

among them to create a broadly applicable relation or model. In the case of twinning in AZ31 the goal is to find a relationship between considered microstructural attributes and twinning events. These results will confirm whether all important aspects have been incorporated into current Mg twinning models, and what may be missing from current models.

Machine Learning has already been utilized for previous research in materials science and other fields to develop constitutive relations that establish structure property relationships (Reddy, Rao et al. 2005, Altinkok and Koker 2006, Xu, Wencong et al. 2006, Koker, Altinkok et al. 2007, Haj-Ali, Kim et al. 2008, Yassar, AbuOmar et al. 2010, Pérez-Benitez and Padovese 2011). However, none of these studies venture to explain the meaning of the models or relationships, but rather treat the machine learning models as a purely black box approach. The goal of the fourth chapter of this dissertation is to create a machine learning model that describes twinning phenomena in Mg, and then use the model to elucidate the physics associated with these events.

One reason that machine learning has been successfully used to solve questions in materials science stems from its ability to sort through a myriad of inputs, such as microstructural attributes, to find complex correlations that are not easily captured by classic techniques. These correlations can be found through any number of available processes that can be broken down into three categories of algorithms: knowledge-based, rule-based, and skill-based (Rasmussen 1983). Knowledge-based learning is equivalent to *ab-initio* studies and therefore requires a greater prior understanding of the studied phenomenon. However, rule-based models provide less structured connections while still maintaining a high level of accuracy. Finally, skill-based algorithms can be compared to complex curve fitting in which the resultant model provides an easily implementable mathematical equation but may have reduced physical

significance. For the case of twinning in AZ31, a rule-based method is most suitable, given that there is not enough information to develop a knowledge-based model, and skill-based models wouldn't provide physically interpretable insights into the causes of twinning.

One common rule-based technique that has been successfully used in previous studies to find structure property relations relies upon decision trees. This method categorically partitions information to maximize the information gain at each level, where information gain is mathematically defined. The result is a hierarchy of attributes (structure or other field property) based divisions that provide explanations or insights into the potential cause of the studied event (see Figure 1-1 for a typical decision tree).

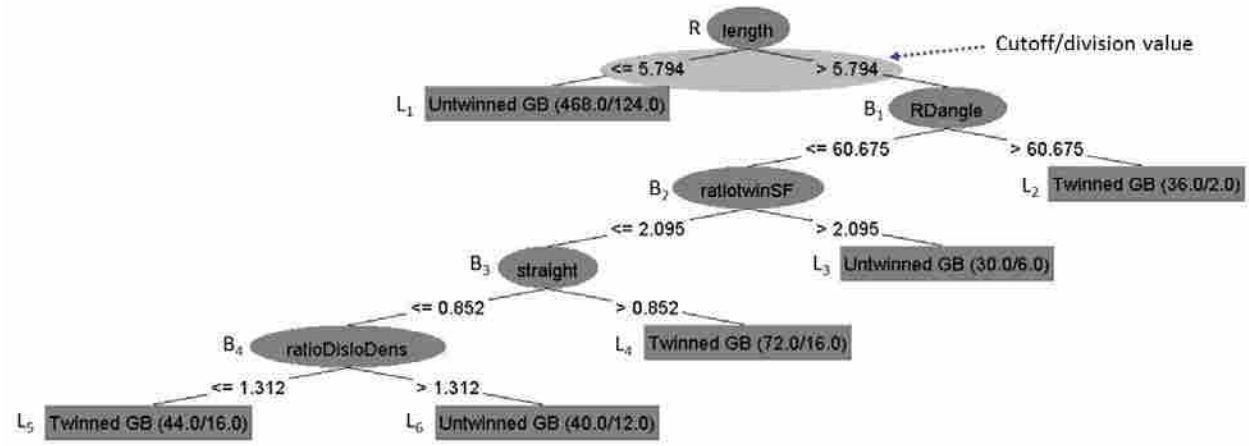


Figure 1-1: A typical decision tree and the associated hierarchy (R: Root, B: Branch, and L: Leaf) present in a decision tree. Note that the values inside the leaves appear in the following format: (# of classified instances / incorrect predictions).

By definition, attributes found nearer to the beginning of the tree, or "root", provide greater information gain than subsequent attributes and are thus considered to be more consequential. For this purpose smaller decision trees are desired since they only retain attributes that have the greatest impact on a particular phenomenon such as deformation twinning in Mg.

With only the most pertinent microstructural features the resultant model becomes more physically understandable.

The ability of decision trees to elucidate the physics behind a particular process, through easily interpretable results, is among the greatest benefits of this machine learning approach. The model created by a decision tree may be interpreted or used in several ways: i) as a model/constitutive relation for the studied event, ii) as a set of insights into the causes of the event that help focus further research, or iii) as a way to find/capture more events for further study. The latter choice provides a framework to systematically increase the data set from which to draw conclusions and thus refine the machine-learning model. In the case of time-consuming data collection procedures, the potential for a refined data collection method is desirable as it has the ability to intelligently guide data collection towards areas of interest. Refined data collection is an area of ongoing research and beyond the scope of this dissertation.

2 QUANTIFYING SURFACE DEFECTS AND TOPOGRAPHY FROM RAW EBSD IMAGES

2.1 Background

The phenomenon of deformation twinning often causes localized changes in morphology which are measureable as surface topography (Stoudt & Hubbard, 2009; Valkonen, 1987). Measuring and quantifying twinning becomes especially important when they provide sites for crack initiation and/or propagation (Barnett, 2007). Deformation twinning, however, is defined by crystallography and not morphology (Christian & Mahajan, 1995). This combination of morphological and crystallographic effects caused by twinning would typically require two measurement tools which would then need some degree of alignment in order to align the two sets of data. Data correlation, or correlative microscopy, can be difficult and resource intensive due to all of the potential differences between the two measurements' geometries. This paper presents a method whereby one technique, electron backscatter diffraction (EBSD), can be used to extract both crystallographic and surface topography of crystalline materials.

Current measurements of micro-scale surface structures can already be measured using a number of techniques including confocal microscopy, laser interferometry, and atomic force microscopy (AFM) for correlation with crystallographic information (Behm, Funke, & Möller, 2013; Vaudin, Stan, Gerbig, & Cook, 2011; Wolfer et al., 2009). Each of these surface techniques usually operates in a separate system from any crystal orientation measurement tool. This leads to the need for an additional technique to acquire the crystal orientations of a sample.

The study of a materials crystallography can be done in several ways, but over the last 20 years electron backscatter diffraction (EBSD) in a scanning electron microscope (SEM) has become a mainstream tool for capturing crystallographic data at the micro- and nano-scales (Wright, 1993). EBSD data is collected when the electron beam of an SEM diffracts off a crystalline point in a sample forming a pattern of diffraction bands, or Kikuchi bands, on a nearby detector. That pattern of bands can then be correlated to a particular crystal orientation. Many recent developments in EBSD relate to extracting information from the images collected by EBSD (Adams, 1997; Sorensen et al., 2014; Wright et al., 2015; Schwartz, 2009; Sitzman et al., 2010). Of these new techniques some can be used to measure surface topography either qualitatively or only under very specific circumstances (Vaudin et al., 2011; Wright et al., 2015).

In order to quantify the topography of a surface with EBSD images under the widest range of samples it is important to understand the mechanisms behind EBSD image formation. As previously mentioned a standard EBSD pattern forms when an electron beam diffracts off of a highly tilted (usually 70^{o}) crystalline sample. More specifically, when the electron beam of the SEM impacts the sample being studied electrons are scattered inside a small volume of the sample. Within this volume electrons experience different levels of energy loss as they interact with the sample material. The electrons that experience very little energy loss are considered backscattered electrons (BSEs). Some of these BSEs diffract off the crystal planes in the sample in such a way that bands of high intensity, or Kikuchi bands, form on the EBSD detector. However, the BSEs that reach the EBSD detector without diffraction information contribute to an additional signal, or background, which will depend on the backscatter coefficient of the sample point and its surface normal. Typically image processing is applied to the EBSD patterns to remove the background such that the diffraction signal is more easily and clearly detected, but

for this study the background signal will be kept and analyzed to provide surface structure relating to topography (see Figure 2-1). In particular, the shape of the EBSD background will be used to extract the surface normal at each point in order to reconstruct the topography.

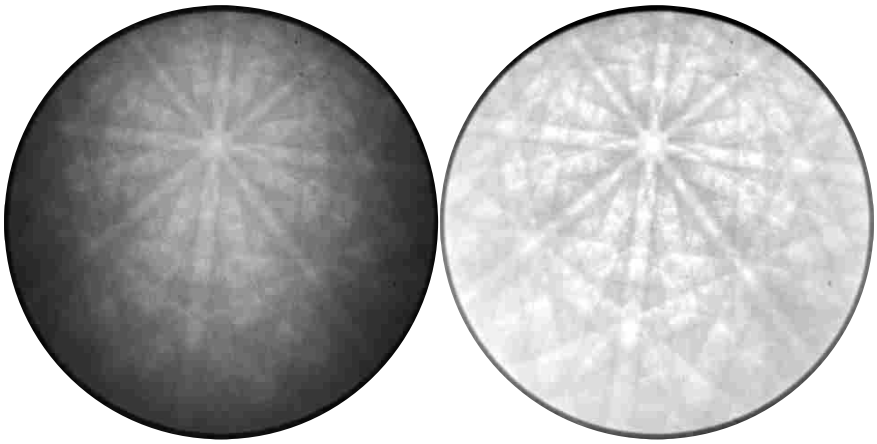


Figure 2-1: Raw EBSD image (left) and background removed EBSD image (right).

2.2 Materials and Methods

The basis for recovering surface topography in this paper is related to the fact that the region of highest background intensity in an EBSD image is governed by the sample/detector geometry and surface normal of the material at each sampled point. The effects of sample tilt on EBSD patterns were demonstrated over fifty years ago by Alam in order to optimize sample detector geometries for collecting EBSD patterns (Alam et al., 1954). In more recent studies Deal et al simulated EBSD backgrounds using Monte Carlo in an effort to find a sample/detector geometry with the most intense part of the background at the same location as the detector x*,y* calibration (Deal et al., 2005). Field also utilized this phenomenon to study faceted surfaces, and qualitatively correlated the facet tilt to image quality (Field, 1997; Wright & Nowell, 2006). The correlations in that research allowed the detection of up to five degrees of inclination. To

accommodate a wider range of surface inclinations this study offers a model based more on the physics of the electron interactions with a sample to determine the surface normal at each point in a scan.

As previously mentioned the surface normal will be measured from the shape of the EBSD background. A typical back can be described in three dimensions (i.e. an intensity at position (x,y) on the EBSD detector). However, for this study the shape of the background will be somewhat simplified by only looking at points that fall within a range of intensities. By applying this threshold to an EBSD image it then appears as an elliptical shape which can be fitted to an equation which parametrically defines an ellipse (Prasad et al, 2013):

$$\frac{\left((x-x_c)\cos\theta+(y-y_c)\sin\theta\right)^2}{a^2}+\frac{\left((x-x_c)\sin\theta-(y-y_c)\cos\theta\right)^2}{b^2}=1 \quad (2\text{-}1)$$

In this equation (x_c,y_c) represents the (x,y) center of the ellipse relative to the center of the EBSD detector, θ is the rotation of the major axis measured from the horizontal direction of the detector, and the major/minor axes are a and b (a is the major axis when a>b and b is the major axis when b>a). The ellipse parameters that were used in this study consist of the (x_c,y_c) center position on the detector, the rotation of the major axis relative to the y-direction, and the lengths of the major and minor axes. The process for extracting the background/ellipse information is illustrated in Figure 2-2.

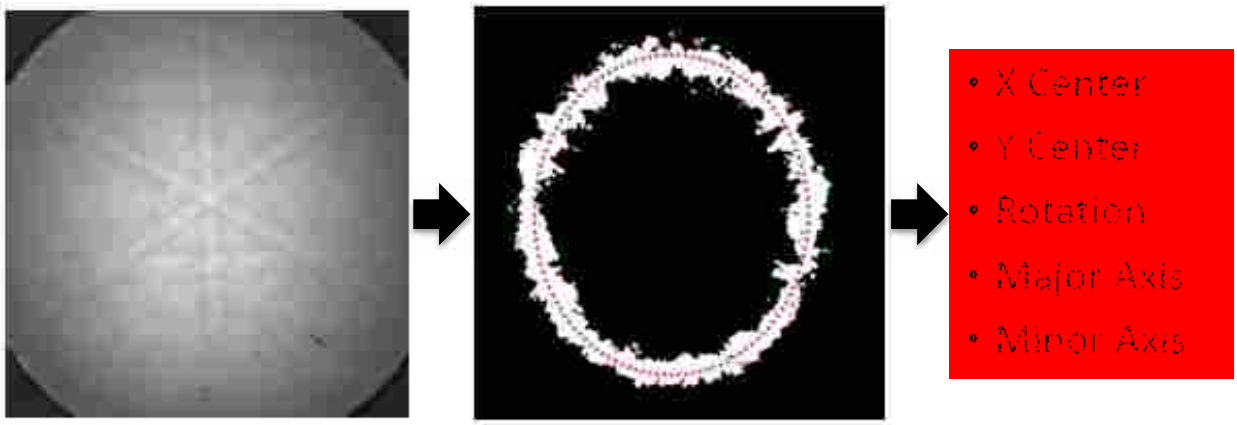


Figure 2-2: Process for extracting background ellipse shape information. First an unprocessed EBSD pattern (left) is thresholded to find the general ellipse shape (middle). The ellipse is then fit by an equation to extract relevant parameters (right).

It is worth noting that the threshold applied to show the ellipse was always chosen so that the horizontal axis was approximately half the width of the entire image. This was done so that the diffraction band contrast was weak enough as to not significantly affect the fitting of the ellipse to the background. Additionally, the resolution of the parameters taken from the ellipse equation depend on the quality of the collected image; in other words, the signal to noise ratio and camera pixel resolution will determine the precision with which the critical ellipse parameters can be determined. For this study the process of data collection was kept consistent to keep signal to noise ratios the same as well as camera pixel resolution.

To extract the surface normal information the ellipse data at each point was coupled with the sample/detector geometry as well as the sample position from which the data came and the particular material properties. The material information was captured through calibration which looked at the average ellipse center over the sample area. The remaining data was stored by default in the EBSD data files. All of this data was combined to extract the surface normal at a point in two pieces: x-tilt and y-tilt. The x-tilt is defined by rotation about the sample x-axis and

y-tilt is defined by rotation about the sample y-axis. Figure 2-3 shows the sample coordinates used.

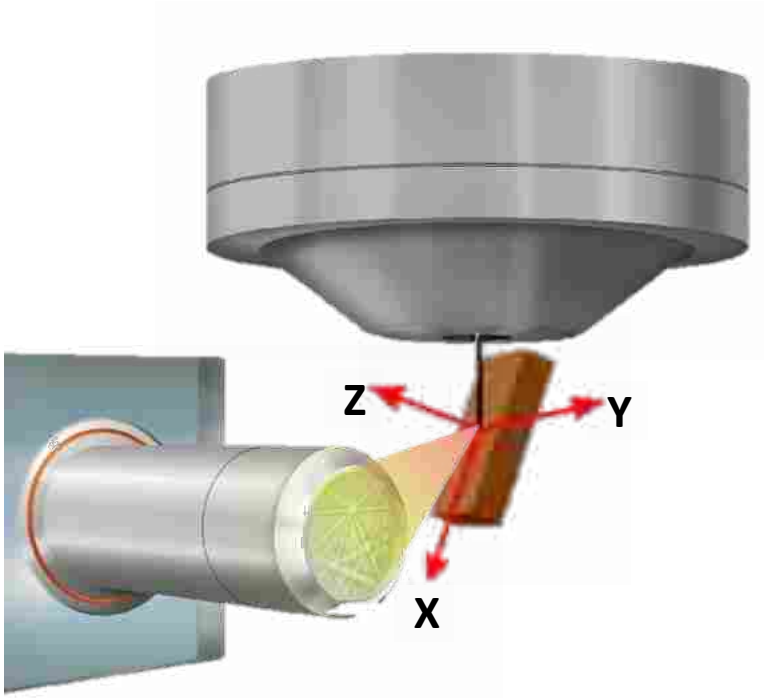


Figure 2-3: SEM/sample/EBSD detector geometry and coordinates.

X-tilt and Y-tilt could either be formed from a series of equations based on the geometry of the data collection or empirically based on samples at known tilt angles. The latter method is used in this study as there is some shift in the background position based on material and the energy of the electron beam. Initial calibration curves were collected on a flat Ge wafer over a relatively small area (100µm x 100µm). Data was collected every 5 degrees about the X-axis ranging from 80 degrees to 60 degrees where standard EBSD occurs at 70 degrees. In the Y-axis data was collected every 5 degrees from -10 degrees to 10 degrees. A wider range of angles was simulated in Matlab based on the geometry of the sample/detector geometry (see Appendix A) simple reflection to confirm the general trends observed by actual measurement. The test results

from the Matlab simulations can be seen in Figure 2-4. The actual measured calibration used for rebuilding the surface topography is shown in Figure 2-5.

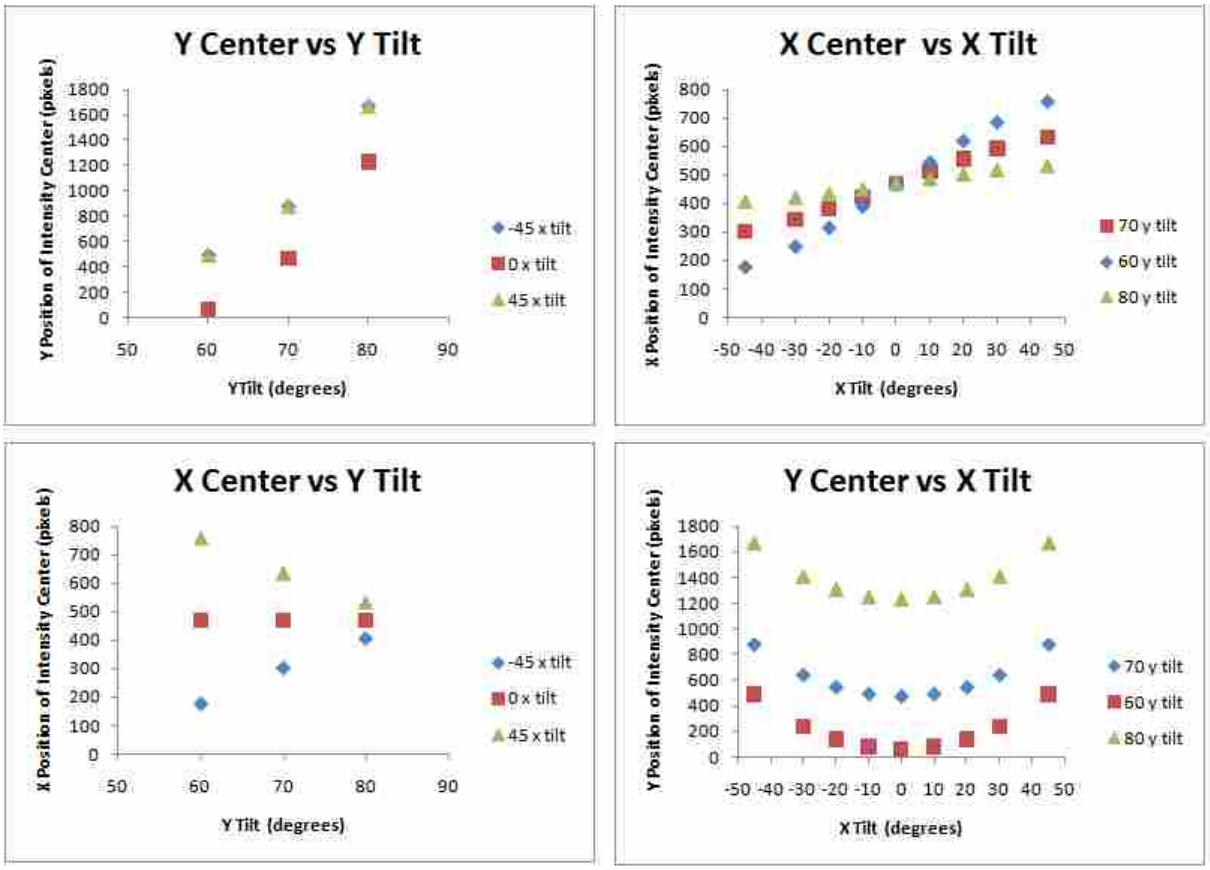


Figure 2-4: Simulated (Appendix A) effects of X or Y tilt on both X or Y ellipse center plotted for three different settings. General effects of sample tilt are mostly linear except for the effect of X tilt on Y center. Clear trends can be seen that demonstrate the nearly linear interaction relationship between X tilt and Y tilt.

In addition to the effects of surface normal on the shape and center of the background ellipse it was observed that the x,y position on the sample influenced the ellipse parameters. This effect arises from changes in the incident angle of the electron beam with the sample as the beam is scanned across a sample. However, these effects are only seen for larger areas (>~500 µm x ~500 µm) which result in the electron beam changing exit angle by approximately one degree.

These additional effects were still ignored on the large area scan due to the effect in the Y direction being so little.

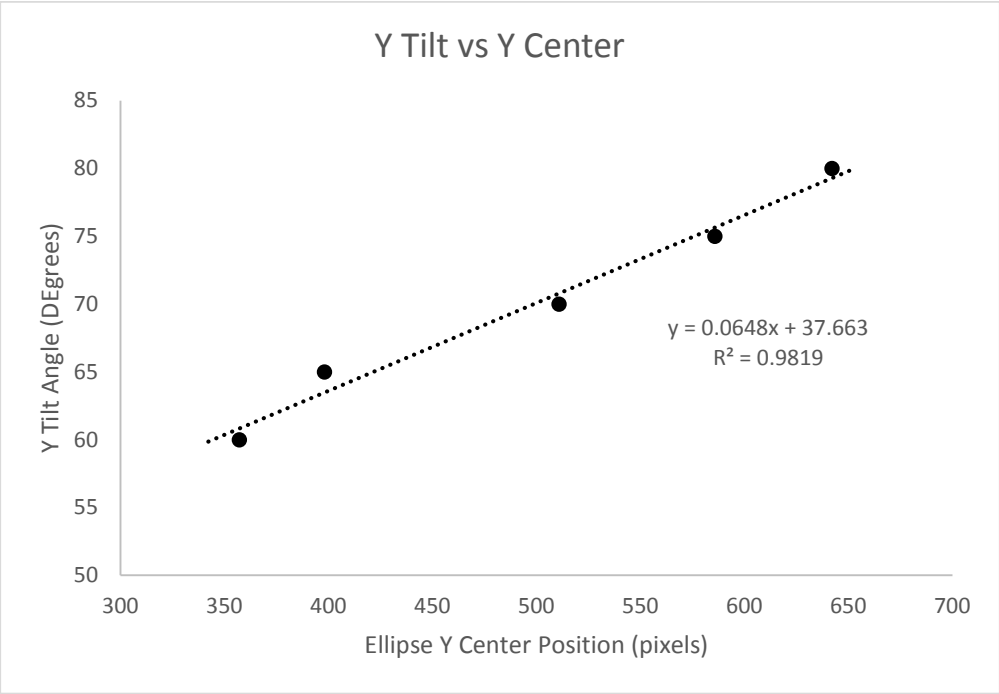


Figure 2-5: Calibration of Y tilt to the background ellipse Y center. This calibration was used for surface reconstruction.

With the calibration data from Figure 2-5, the Y-tilt at each point was calculated. Assuming a starting edge height of zero the surface topography was reconstructed starting at the flat edge and iteratively changing sample height based on the following formula:

$$\boldsymbol{h_i = h_{i-1} + dsin\theta} \tag{2-2}$$

where h_i is the height of the sample at location i for each row of reconstruction. In this formula d is the spacing between data points, or step size, and θ represents angle of tilt. In the case of Y-tilt

θ was adjusted by 70 degrees to account for the overall Y-tilt applied to the sample to perform EBSD.

This study will also use EBSD images to analyze cracks/surface voids in a sample, which produce a different effect from electron beam interactions than normal surface topography. As opposed to changing the direction of electron reflection cracks and voids block the reflection of the electron beam relative to the EBSD detector. Therefore for this particular type of surface topography the shadowing that occurs on the EBSD pattern when the electron beam encounters a surface void will be analyzed.

Similar to the method of surface topography the EBSD images will be collected in raw form and passed through a threshold to highlight dark regions in the pattern which will be called shadowing. The threshold was placed at 30% of the maximum possible intensity meaning anything below this threshold was considered for shadowing analysis. The resulting image was passed through a convolution mask to highlight grouping of shadows. If the resultant convolution surface at a point was above a certain value then that point was considered a defect. This process is illustrated in Figure 2-6.

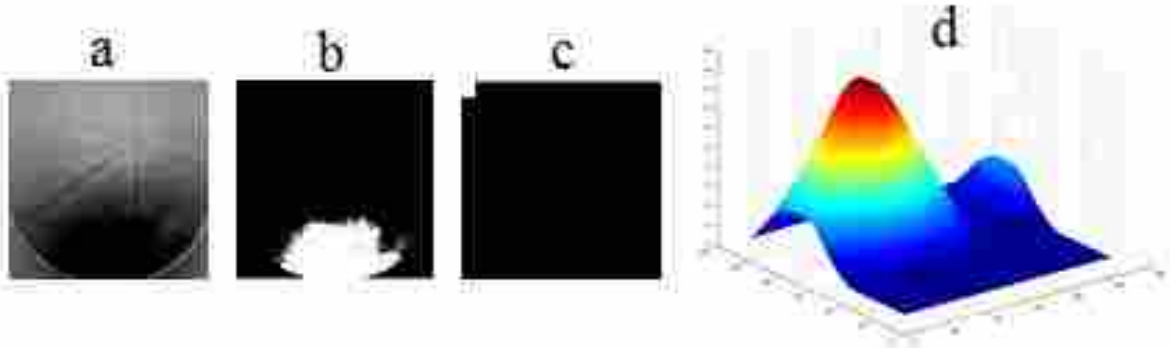


Figure 2-6: Procedure for image shadowing due to cracks. EBSD images are passed through a threshold process and the resultant images are then convolved with the mask (c). After convolution, if the resultant image peak was above a specified threshold then the point was deemed a crack.

Scanned cracks were tested in one of two positions: horizontal and 45° relative to the sample x-y coordinates. Cracks which were scanned ranged in widths from 0.5 microns to 6.7 microns which represents a sufficient sampling to understand the effects of cracks on EBSD imaging. In order to test the resolution of this method scans were taken at several step sizes from 0.25 microns to 11 microns.

In addition to identifying the presence of surface cracks this technique can be used to determine the orientation of a crack. This is performed by locating the centroid of the shadow relative to the center of the EBSD detector (see Figure ##). Crack orientation is then defined as the line perpendicular to the segment formed between the center of the detector and shadow centroid. Prior to acquiring EBSD images of cracked regions optical measurements were taken in the SEM to obtain width and shape of several cracks. These were used as a guide to find the range of crack width detectability.

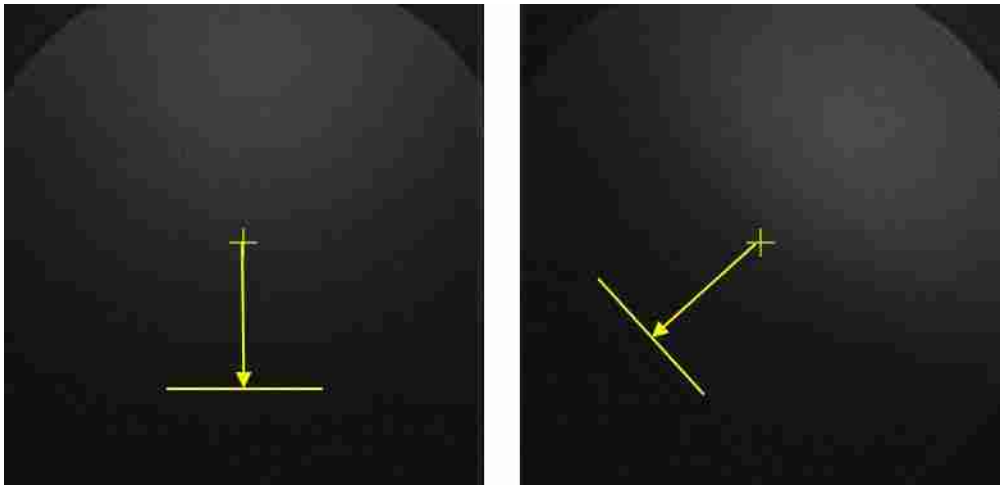


Figure 2-7: EBSD images demonstrating the effect of crack direction on shadowing on the EBSD detector. Both a horizontal crack (left) and crack rotated 45º (right) are shown. The arrow from the center of the detector to the centroid of the shadow is perpendicular to the crack direction.

Data for this research was collected using two scanning electron microscopes, Philips XL30 S-FEG and FEI Helios Nanolab 600, at an accelerating voltage of 20 kV. Samples analyzed for surface topography included a US nickel which were used as proof of concept and to demonstrate capabilities over large areas. Additionally, a small 3 x 3 μm^2 region of a twinned Ta specimen was analyzed. The data collected for crack/void detection came from regions of an ultrasonically consolidated Ni sample with areas of incomplete joining. EDAX/TSL orientation imaging (OIM) software was used for EBSD data collection. EBSD images were saved at a resolution of 936x936 or 468x468 pixels corresponding to an image diameter of approximately 32 mm. Samples in the microscope had five available degrees of freedom (Figure 2-3): x-y-z motion, z-rotation and y-tilt. It should also be noted that with the current setup, y-tilt was set manually to a precision of 0.1°. X-tilt in the microscope is constant and therefore any variation in x-tilt was implemented via special sample holders.

2.3 Surface Topography Measurements

The letter "O" was scanned using a 10 micron step size without any polishing resulting in weak or no Kikuchi bands in the EBSD images. Even without Kikuchi bands the method for reconstructing the coin surface still proved effective. The reconstruction of the "O" is shown in Figure 2-7 along with the data collected from the same area by a stylus profilometer. While results for the reconstruction of the coin did not yield exact results over the entire surface they did demonstrate reasonably accurate height measurements from the reconstruction method described over a 1 mm^2 area.

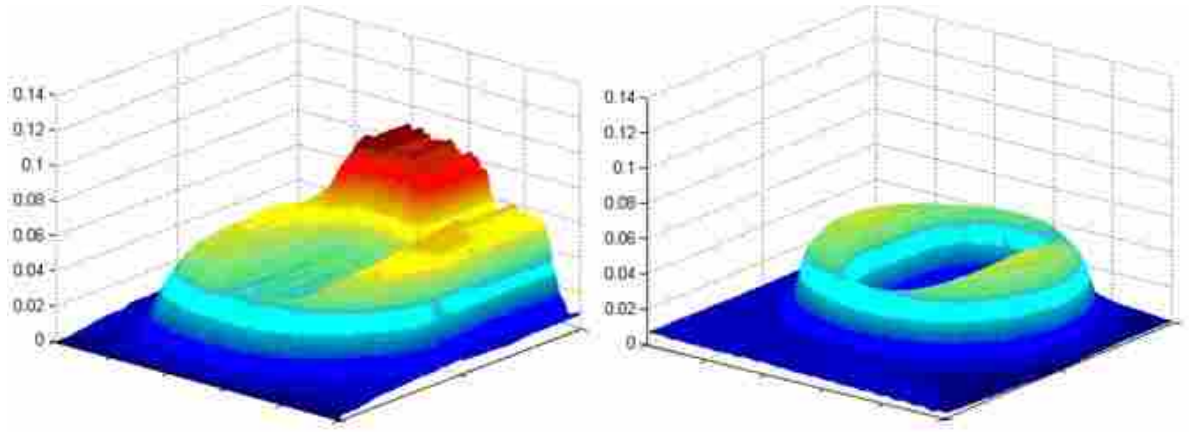


Figure 2-8: Angled view of (left) the reconstructed surface vs. (right) measured profilometer surface using the same color scale for both (neither surface is drawn to scale). The height profiles match fairly well on the front edge of the coin. Back end errors are due to the inability of the microscope to capture angles beyond 85 degree tilt.

The most visible measurement errors occurred due to the sample geometry which causes the electron beam to not reach points on the surface. As might be expected, when the electron beam doesn't reach the intended point no information can be collected. Such is the case when the surface normal is around 90 degrees or greater which corresponds to about a 20 degree inclination limitation in one direction. The maximum inclination seen in the "O" was around 40 degrees which exceeds the 20 degree threshold. The observed errors from this effect effectively cause points to be skipped by the electron beam and thus their contribution to inclination goes unaccounted. X-tilt values of ±40° were present in the coin and were detected without problem meaning the X-tilt range was valid up to ±40°.

The secondary issue seen in the reconstruction appears as streaking in the direction of reconstruction. As shown in Fig. 2-7 the reconstruction starts at an edge, assumed to be straight, and is built up in rows. If the beginning edge is not straight then the reconstruction can start off incorrectly. Alternatively, the step size chosen for this particular scan may have been too large as to cause certain features to be missed or highlighted which could cause sudden changes in the

reconstructed height which would then affect every subsequent point in a row of reconstruction. The step size used for this scan was 10 μm which is much larger than the estimated probe size of 100 nm, so even with the scanned area there are many areas from which no information is acquired.

However, in spite of many of the errors, the measured peak height of the coin matches within ±1 micron of the measurements taken via a stylus profilometer (12.5 μm radius tip). This represents a ~2% error in height measurement at the peak which was 500 μm from the starting edge of reconstruction. In addition to accurate height measurement the general shape of the "O" was captured along with many local features such as scratches.

With the "O" serving as a proof of concept for surface reconstruction, a small nearly 3 μm x 3 μm area of twinned Ta was also studied (Figure 2-8). In rebuilding the Ta surface from EBSD image data the topography was seen to correlate to certain microstructural features (Raabe, 2003; Becker, 1998). Figure 2-8 shows several maps of the microstructure derived from different metrics. Of all these of these maps the topography seems to be best aligned with the Grain Reference Orientation Deviation – Angle map. This map shows changes in orientation within a grain by comparing the orientation at each point in a grain to some reference orientation (usually the average grain orientation).

In addition to the topography aligning with certain measureable microstructural features the height increase caused by the twin was 94.9 nm which was only slightly greater than the step size used to collect this data. The area scanned was too small to be marked with the equipment used so the area could not be found for measurement with another surface technique; however,

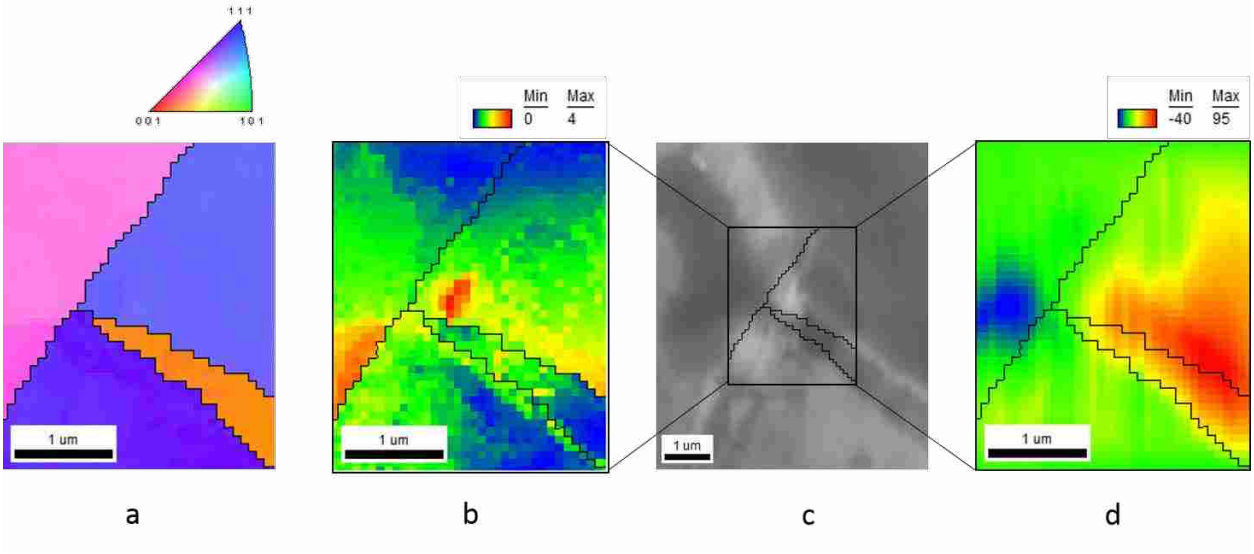


Figure 2-9: Inverse pole figure (a), grain reference orientation deviation - angle map in degrees (b), secondary electron image (c), and reconstructed surface in nm (d). All images have the twin boundaries overlaid to facilitate visual correlation. The maximum height of the surface appears to line up with one of the twin/parent interfaces while the low point seems to occur at an area of local orientation change within the grain on the top-left.

based on the orientation, size, and type of twin a range of theoretical heights was calculated. The twin plane was a {321} plane and the <1$\bar{4}$5> direction. If the twin direction is assumed to slip and cause topography changes then a height of 124.7 nm can be expected. However, if twinning dislocations of $\frac{1}{6}[\bar{1}11]$ are assumed to cause twin topography then the maximum height change normal to the sample surface would be 84.6 nm (Christian, 1995). Therefore, the reconstructed peak height of 94.9 nm is only 24% less than the 124.7 nm estimate and 12% greater than the 84.6 nm estimate. In either case the reconstructed height was within 30 nm of the theoretical height.

In addition to reconstructing surface topography, the parameters associated with the background ellipse provide other insights into the surface structure. Figure 2-9 shows the ellipse x-center, y-center, and area parameters of the “O” which each offer a unique view of the same

surface structure. The ability to separate features of surface structure provides additional opportunities when studying a given material.

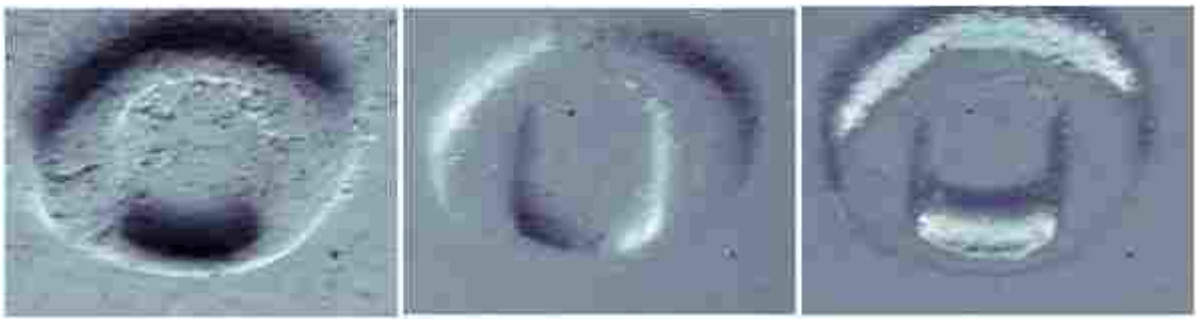


Figure 2-10: Maps of y-center (left), x-center (middle), and area (right) from the letter "O" on a coin. Each map reveals a different characteristic of the coin. The roughness of the sample is apparent in the y-center image due to larger variability in y-center than x-center.

2.4 Crack Detection

EBSD scans of the UC nickel defects/cracks were all successfully analyzed for the presence of the cracks (see Figure 2-10). Figure 2-11 shows an example of a cracked region with successful identification and mapping of the crack. Additionally, the crack direction was accurately measured (±3^{o}) in the horizontal and 45 degree orientations based on the use of the location of the crack-induced shadow centroid. Step size had a clear effect on the detection limits of this method; larger step sizes were less successful at capturing cracks within the grid of scanning when they exceeded the crack width (see Figure 2-12). For example, the 6.7 μm and 1 μm crack are completely missed by step sizes larger than the crack width. In Figure 2-13 the data for each step size was fit to a Gaussian shape which helps visualize the spatial uncertainty caused by using larger step sizes. Errors from larger step sizes are not certain, but can happen depending on how the scan steps line up with the crack.

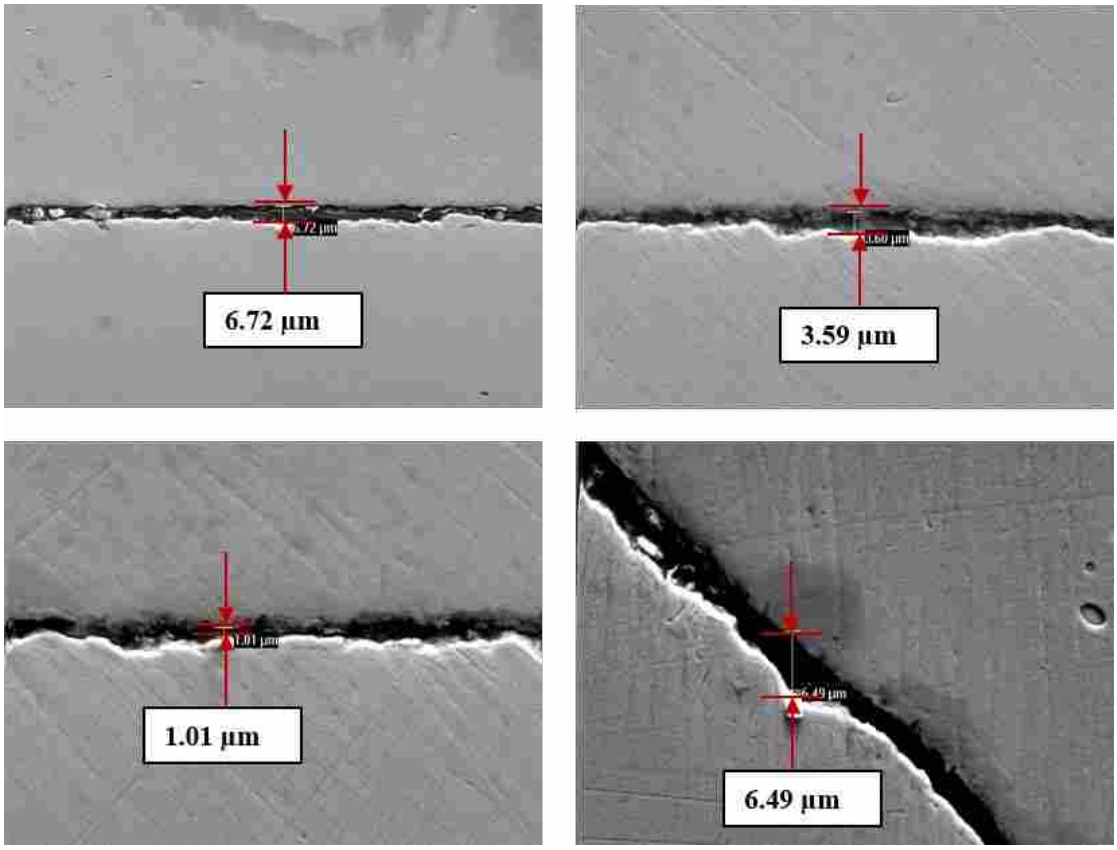


Figure 2-11: SEM images of several observed cracks. 6.7 µm (top left), 3.6 µm (to right), 1.0 µm (bottom left), and 6.5 µm at 45° (bottom right).

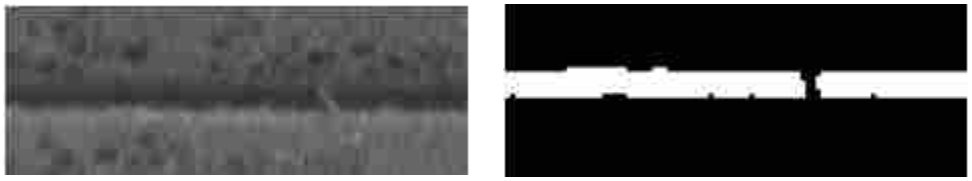


Figure 2-12: (left) SEM image of UC nickel crack with (right) reconstructed defect map. The resolution of reconstruction is a function of the step size of the scan.

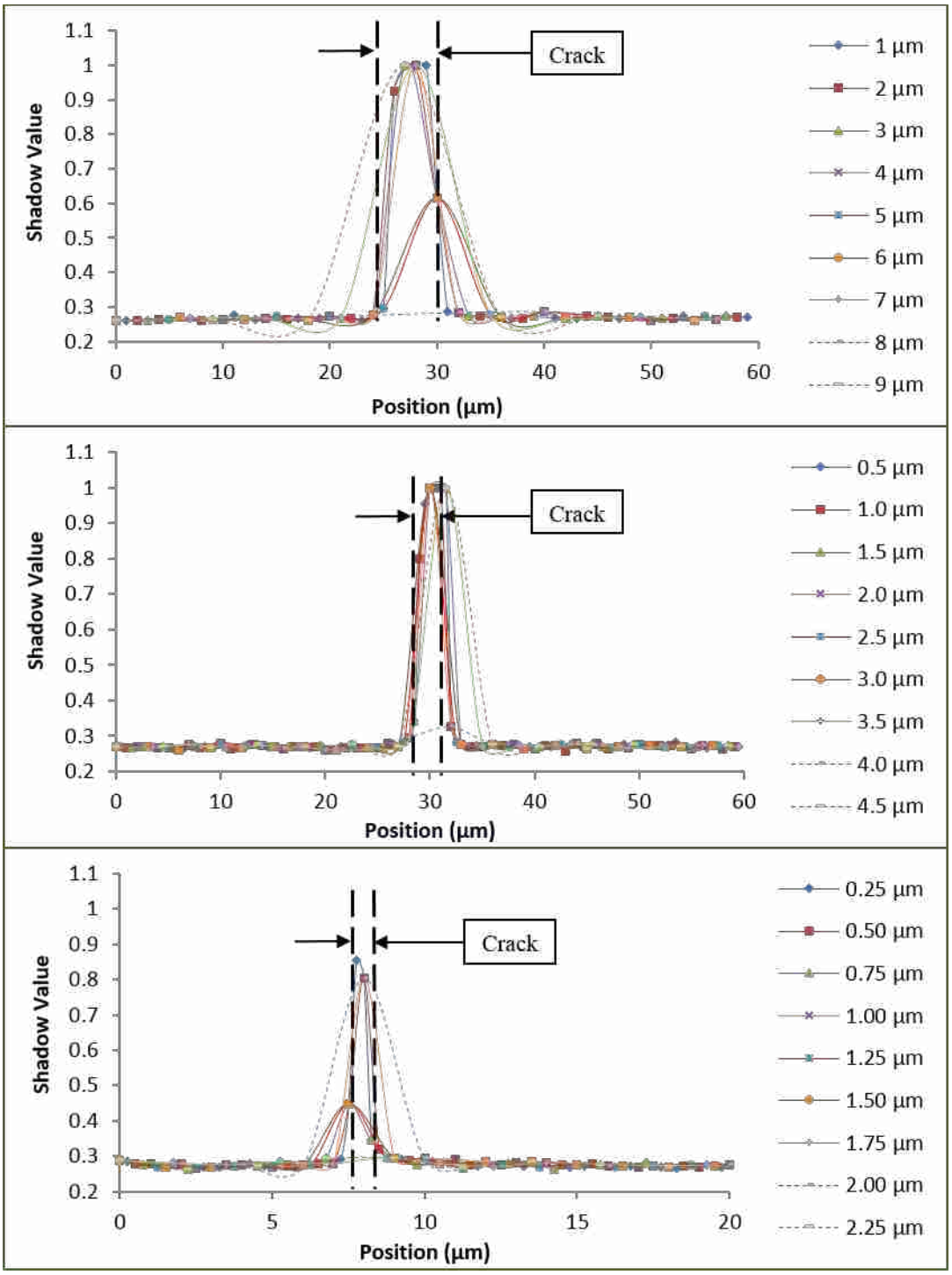


Figure 2-13: Shadow values across several cracks at varying scan step sizes. 6.7 µm crack (top), 3.6 µm (middle), and 1.0 µm (bottom). The actual width of the scanned crack is denoted by the dashed black lines. The curves are a Gaussian fit to collected data points which demonstrate greater spatial uncertainty for larger step sizes. This uncertainty can make a crack appear larger than reality.

While all of the cracks in this study were detectable with the smallest step sizes used (250 nm) it is often desirable to study cracks close to their initiation. The minimum crack size that this method could detect was calculated by measuring the convolution surface peak height for various crack widths and extrapolating back to a shadow value of zero. This information is shown in Figure 2-13. The data was found to fit a logarithmic curve up to a maximum value around crack widths of 3.5 µm. It is reasoned that this maximum value is tied to the area that could receive signal, or detector area. Knowing this, cracks greater than 3.5 µm were excluded from calculating the logarithmic fit seen in Figure 2-14.

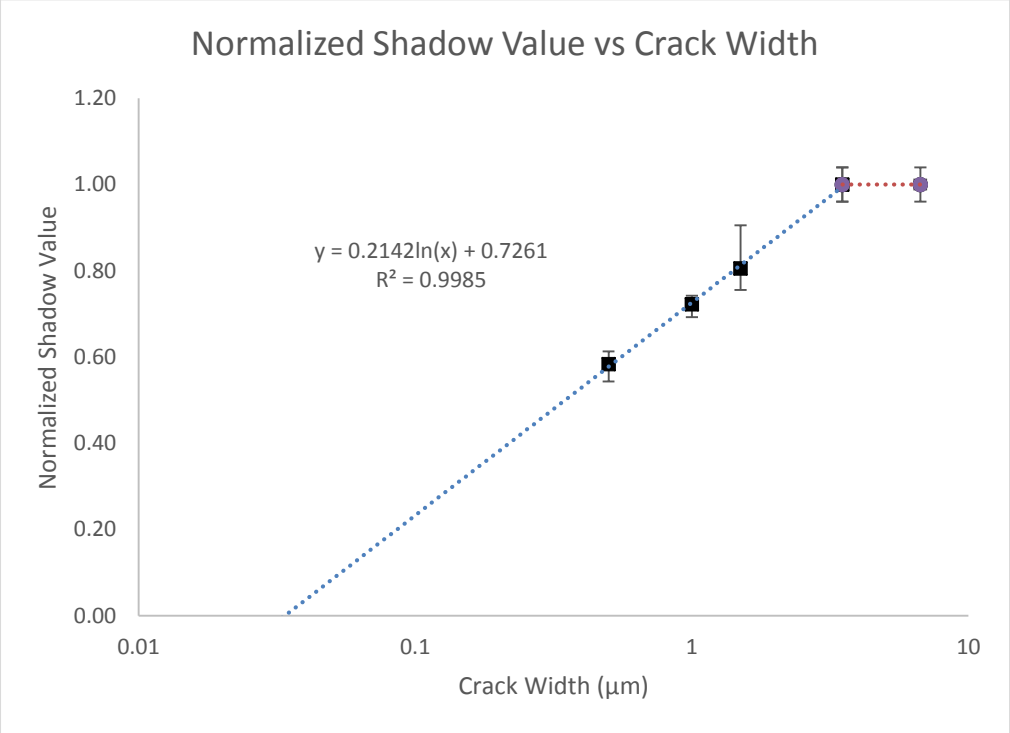


Figure 2-14: Normalized shadow value measurements for several crack widths. The x-axis is logarithmic with the data reaching a maximum shadow level for cracks greater than or equal to 3.5 µm.

Using the calculated formula for the relationship between crack width and shadow value the smallest resolvable crack could be calculated by extrapolation. Based on this extrapolation

the limiting size of crack that can be studied using the method in this paper is ~34 nm. This gives some indication of resolvable cracks, but the true resolution of this method is uncertain due to the extrapolation stretching over an order of magnitude.

2.5 Summary

Understanding the physics of electron diffraction as it pertains to collection of EBSD data allows for more information to be extracted from an EBSD image. This study has used a knowledge of backscattered electrons that do not have diffraction information to quantify surface topography and cracks/defects. Because they are collected with an EBSD detector both of these new tools could be used in conjunction with standard EBSD data where no data alignment would be required.

Reconstruction of surface topography was achieved by measuring the surface normal at a point which was found to be related to certain parameters of the background ellipse shape. Reconstructed surfaces consisted of a stamped letter "O" on a US nickel which was nearly 1 x 1 mm^2 and a much smaller area, approximately 3 x 3 μm^2, of twinned tantalum. The reconstruction of the "O" demonstrated the ability to accurately quantify surface topography even over large areas (2%, or 1 μm, error over 500 μm). This particular reconstruction also showed a limitation in the Y-axis rotation where tilts >~85º where the electron beam was nearly parallel to the sample. Analysis of the twinned tantalum surface also revealed the method's ability to perform reconstruction of small-scale structures to within tens of nanometers which could then be correlated to other orientation data collected from the same EBSD detector.

EBSD images were also used to find cracks/defects in a UC Nickel sample based on the extent and location of shadowing in an image. Shadowing in the EBSD image is caused by electrons being blocked from reaching the detector. Results from this study indicate that the

method is viable for both detecting a defect and determining its orientation. However, searching for cracks with this method might require some level of prior knowledge about the size of cracks or defects which are present in order to select an appropriate step size for scanning Measurements of several crack widths were used to extrapolate for the minimum theoretical crack width that could be measured; this was estimated to be 34 nm.

3 IMPROVED DETECTION AND SPATIAL RESOLUTION OF TWINS VIA TRACKING OF INDIVIDUAL KIKUCHI BAND INTENSITY OF EBSD PATTERNS: APPLICATIONS TO MG AZ31

3.1 Background

With the ever-increasing interest in lightweight metals to improve automotive efficiency, Magnesium has received a lot of research attention. One of the critical findings deals with the role of deformation twinning in limiting the formability of Mg alloys, such as AZ31 (Klimanek and Pötzsch 2002, Barnett 2007, Beyerlein, McCabe et al. 2011, Scott, Miles et al. 2013, Barnett 2007). In AZ31 there are two basic twin modes: compression twinning and tension twinning (Barnett 2007, Barnett 2007). The latter phenomenon forms fairly large, easy to detect twinned regions within parent grains, whereas the former tends to form extremely thin twins that are on the order of 100 nm. Furthermore, combined modes exist in the form of double twins that can lead to even finer structures. Moreover it is desirable to measure the onset of twinning as close to the point of initiation as possible in order to gain valuable insights regarding the exact conditions of initiation. While features less than 100 nm are well within the detectable range for a modern scanning electron microscope (SEM), electron backscatter diffraction (EBSD) techniques that identify twins via crystal orientation rely on a much larger spatial resolution (Zaefferer 2007, Zaefferer 2007). This resolution is influenced by two factors: 1. the size of the electron interaction volume (IV), which can be several hundred nanometers for light metals, which means that the structure data presented in the EBSD pattern comes from a relatively broad area; and 2. The grid size for the raster pattern, which is often limited either by time constraints or by such

issues as charging effects for very fine scans. The second factor means that compression twins can be completely missed during a scan by being stepped over. The first factor results in uncertainty in the identification of small twins due to the EBSD pattern containing information from both the twin and its surrounding parent crystal (Sorensen, Basinger et al. 2014). This is the resolution issue that will form the focus of this paper.

Because of this ambiguity in identifying compression twins with standard EBSD, researchers have turned to higher-resolution transmission techniques, either in the transmission electron microscope (TEM) or in the SEM with a technique known as transmission EBSD (tEBSD) (Geiss, Keller et al. 2010). In either case, the sample preparation is difficult and limits the size of an area that can be studied. In order to facilitate the study of these thin compression twins with standard EBSD a new method is required which can provide the ability to identify twinned regions within parent grains, even when the EBSD patterns contain information from both twinned and untwinned regions. This paper proposes a new method to resolve this issue by considering the individual Kikuchi band intensities within EBSD patterns that are taken from the region about a twin. Lattice planes that are common to the twin and parent grain retain a constant band intensity in the region, while other planes show distinct dips in intensity, thus providing a telltale indicator of twin presence.

As previously mentioned, the compression twins seen in AZ31 are extremely thin and can be difficult if not impossible to detect with standard EBSD. While modern SEMs can achieve extremely small beam spot sizes on specimens relatively independent of the material composition the interaction volume (IV) is inversely proportional to the density of a material. This means that a material such as Mg with a density of 1.738 g/cm^3 will have an interaction volume that is hundreds of nanometers across. Ren et al. (Ren, Kenik et al. 1998) found that at

20 keV copper exhibited lateral and longitudinal resolution of ~50 nm and ~150–170 nm, respectively; for nickel the values were ~50 nm and ~200 nm; for aluminum they were ~200 nm and ~750 nm, consistent with the Hjelen and Nes results (Hjelen 1990). Hence, for a 100nm compression twin in Mg, even if the electron beam focuses on the twin the structure information from the twin will often be swamped by the information from the parent grain that fills the majority of the interaction volume. Hence standard EBSD software packages will not identify the twin.

The size of the IV can be reduced by decreasing the accelerating voltage used to collect data. Stefan Zaefferer et al demonstrated this effect on steels (Steinmetz and Zaefferer 2010). However, even with this technique the IV in Mg will be sizeable compared to compression twin dimensions, not to mention slower collection times at lower accelerating voltage.

In order to understand the techniques that will be proposed in this paper it is important to understand some basic principles of EBSD. Historically EBSD is an SEM based technique for measuring and mapping the orientation of crystalline materials (Wright and Adams 1991, Adams, Wright et al. 1993, Wright 1993, Brewer, Field et al. 2009). However, recent advances have also allowed for the measurement of elastic strain and dislocation density (Humphreys 2004, Demir, Raabe et al. 2009, Basinger, Fullwood et al. 2011). Various groups are working on extracting even more detailed information regarding crystal lattice structure from the diffraction patterns (Wilkinson, Meaden et al. 2006, Kacher, Landon et al. 2009, Villert, Maurice et al. 2009, Fullwood, Adams et al. 2014). In order to find additional information contained within EBSD data it is important to understand the method setup. First, a highly polished sample is placed in the SEM at an angle relative to the electron beam, usually at 70 degrees from the

horizontal plane. When the electron beam of the SEM contacts the surface of the sample, the electrons spread out in what has been described as a teardrop shape within the sample (Goldstein, Newbury et al. 2003). Several interactions occur between the incoming electrons and the sample that produce a range of outputs including secondary electrons, backscattered electrons, and x-rays (Joy 1995, Ren, Kenik et al. 1998, Tao and Eades 2004). These interactions occur within the IV previously discussed. Some of the backscattered electrons diffract off crystal planes to create bands of high intensity called Kikuchi bands on a nearby detector – usually a phosphor screen.

In the traditional EBSD approach, the Kikuchi bands are approximated as straight lines, and detected using a Hough transform approach (Fig. 3-1). In Hough space, the bands form local peaks; the relative and absolute positions of these peaks enable bands to be correlated with crystal planes, and allow determination of lattice orientation. In Hough space a line is defined by the following equation:

$$\boldsymbol{\rho = x\cos(\theta) + y\sin(\theta)} \quad (3\text{-}1)$$

where ρ is the perpendicular distance, or line segment, from the center of the image to the line and θ is the angle between the horizontal and perpendicular segment. A typical EBSD pattern with its associated Hough transform is shown in Figure 3-1.

The height of the peaks relative to the background signal gives information regarding the intensity of diffraction of a selected Kikuchi band and the noise in the system, including the presence of grain boundaries and other lattice defects within the interaction volume. A metric

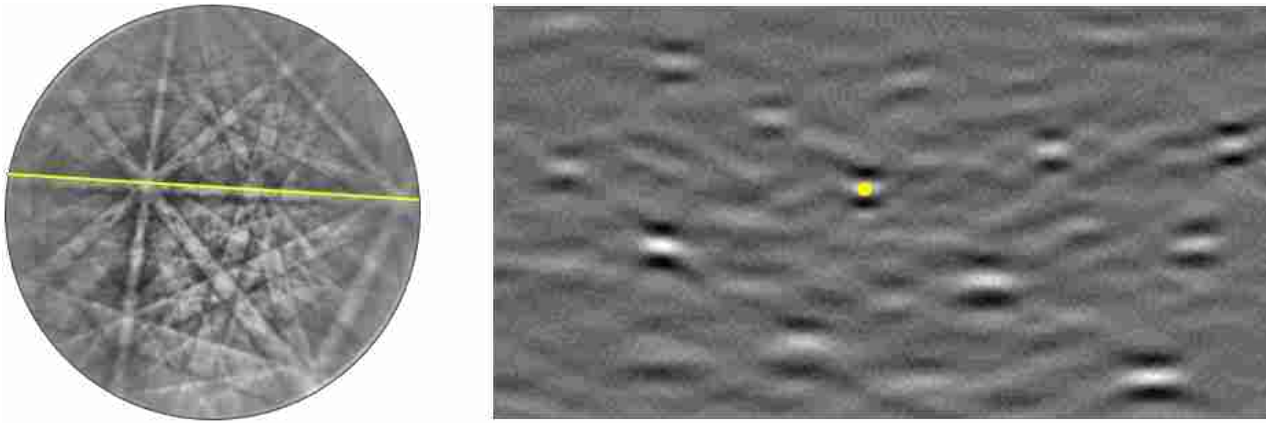


Figure 3-1: Example EBSD pattern (left) and Hough transform of that pattern (right). The yellow line segment on the EBSD pattern corresponds to the yellow dot in the Hough transform. Because Kikuchi bands are not single pixel lines but rather have some width associated with them they appear as local peaks in Hough space as opposed to single points. The highest point of a selected Hough peak is used as the position for the Kikuchi band.

frequently used to quantify the diffraction signal to noise ratio is known as image quality (IQ) (Wright and Nowell 2006). Image quality is calculated with the following equation:

$$\boldsymbol{IQ} = \frac{\mathbf{1}}{\boldsymbol{N}} \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{N}} \boldsymbol{H}(\boldsymbol{\rho_i}, \boldsymbol{\theta_i}) \qquad (3\text{-}2)$$

In this equation, N is the number of diffraction bands being considered (a value that is chosen by experience) and H is the height of the Hough peak, corresponding to a band at location (ρ_i,θ_i).

As seen in Fig 3-2 when scan points approach grain boundaries the IQ drops. For this reason, maps of IQ create a view of the microstructure that provides similar information to optical images of etched microstructures which highlight any boundaries present. Image quality near grain boundaries and other defects is generally low; when the IV overlaps regions of different lattice structure the intensity from each structure is lower. In this case, traditional EBSD software identifies ('indexes') the lattice orientation associated with the stronger signal, and

subsequently ignores the weaker bands that are produced by the neighboring grain (see Fig 3-3). From these observations, it is generally assumed that sharp dips in IQ correlate with grain boundaries or other misorientation boundaries. However, in the case of very thin twins the boundaries are so close together that they appear as streaks in the IQ map which might be confused with cracks or some other defect structure. This idea formed the basis for a morphological twin identification routine developed by Chen et al where regions of low IQ were considered twins when less than two degrees misorientation was detected (Chen, Lui et al. 2009). While the technique developed in that study was helpful for the application being tested it could very well have confused other surface topography (i.e. a scratch) with a twin.

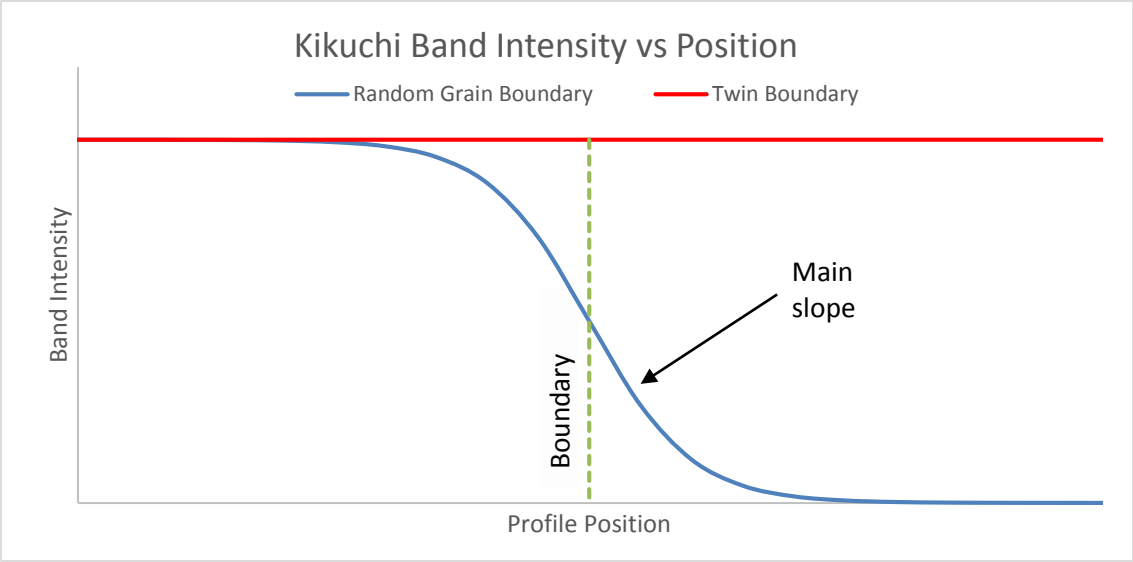


Figure 3-2 - Theoretical change in Kikuchi band intensity across a twin boundary. Note that the twin plane intensity remains constant while the intensity of a band unrelated to twinning drops. The profile of this drop corresponds to the boundary's 3D inclination.

To remove the possibility of topography effects, the basis of the method described in this paper is more crystallographic and relies on the extraction of information from individual Kikuchi bands as they relate to particular crystal planes. This ignores the averaging used to

calculate IQ but still uses the same measurement of Hough peak height to quantify Kikuchi band intensity.

Additionally, we consider only contact twins – twins that share a plane with their parent grain. For this type of twin, the presence of a common plane between the parent and child means that as the IV crosses the twin boundary, the IQ of non-common planes drops, but that of the common plane should remain fairly constant. If the band for the common plane lies on the phosphor, even a very fine compression twin can be identified by the fact that the band IQ for this plane does not dip sharply, while the aggregate IQ does. This effect provides a method whereby very thin twins, such as compression twins found in the Mg alloy AZ31, can be unequivocally identified where standard EBSD methods and morphological approaches only allude to their presence.

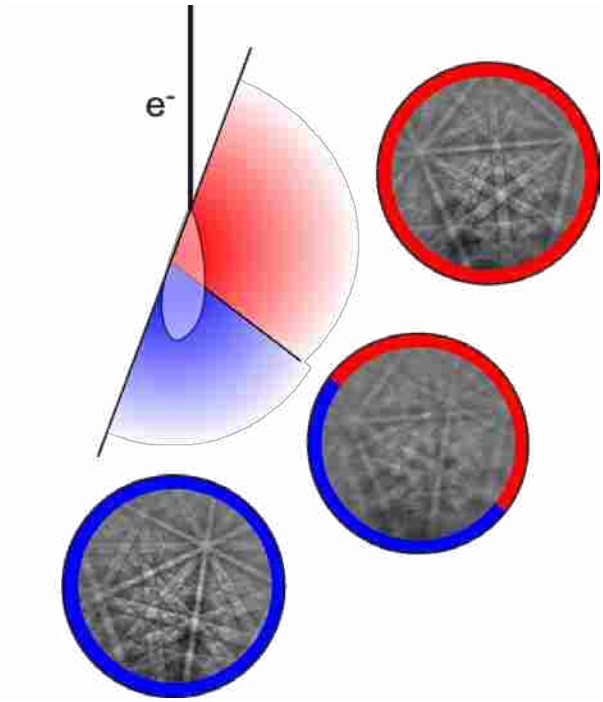


Figure 3-3 - Schematic of mixing patterns at or near a grain boundary. Even if EBSD patterns are sharp on both sides of the grain boundary the pattern near the boundary can appear fairly blurred.

3.2 Materials and Methods

Data for this paper was taken from three different materials: Inconel 600 (Face-Centered Cubic/FCC), Tantalum (Body-Centered Cubic/BCC), and Magnesium alloy AZ31 (Hexagonal Close Packed/HCP). All three of the materials contained contact-twin boundaries as well as other random grain boundaries. Data was collected from several grain boundaries with the focus being on the twin boundaries in order to demonstrate the method presented in this paper. EBSD patterns for the samples were collected at a resolution 480x480 pixels using an EDAX Hikari camera. In addition, the Inconel 600 was also scanned at pixel resolutions of 240x240 and 60x60 in order to determine whether the technique was viable for lower resolution images.

Samples used in this study were prepared via standard EBSD mechanical polishing methods described in Khosravani et al (Khosravani, Scott et al. 2013). They were scanned with an accelerating voltage of 20 keV which was considered to be typical for many EBSD scans. Additionally, the camera gain was minimized to reduce any effects of noise in the EBSD patterns that could degrade image measurements. The Inconel 600, Tantalum, and AZ31 were scanned with step sizes of 100 nm, 20 nm, and 300 nm respectively. The twins in the Inconel and Ta were large enough such that the boundaries and twins were easily definable and thus served as a demonstration of the technique as it would be applied to AZ31.

For twin boundaries, the same drop of IQ occurs as across a random grain boundary; however, by definition twins share at least one crystallographic plane which means that the band intensity of the shared plane, and possibly others, should remain constant across the boundary. With traditional measures of Image Quality that average the intensity of several Kikuchi bands the consistency of the twin plane intensity goes unnoticed. This study looks at the intensity of individual Kikuchi bands across boundaries where image quality is lower than the rest of a grain

in order to identify twin boundaries based on the constant intensity of certain crystal planes in an EBSD pattern. In addition to identifying twin boundaries, this method can be applied to all grain boundaries affected by IQ to extract additional information such as grain boundary inclination. A basic diagram of both concepts is shown in Fig. 3-2.

Theoretically, the flat line would represent the intensity of the twin plane, or some other symmetric plane related to twinning, diffracting onto the EBSD detector. Otherwise, the intensity of an individual crystal plane not tied to the twinning symmetry would drop off as the beam moved from one grain to the next. The rate of drop of the Kikuchi band intensity correlates directly to the inclination of a grain boundary. To determine the exact tilt more detail would be required corresponding to the interaction volume character. For the purposes of this study, exact information about the interaction volume will be ignored and thus grain boundary inclination will not be addressed further.

With traditional measures of IQ, the Kikuchi bands used for quantification change as a scan moves across a grain boundary depending on which bands comprise the dominant, or most intense, signal. For this study however, the observed Kikuchi bands used for intensity calculations were selected from the ten most intense bands at a reference point within each grain. Line profiles of IQ and individual Kikuchi band intensity will be generated. The individual band intensities were measure directly from the Hough transform based on the bands selected from the reference point. In order to clearly identify the bands and their intensities effort was made to match the Hough routine found in OIM DC which has been proven to be fairly robust.

Additionally, each band detected via the Hough transform will be correlated to its respective crystallographic plane. By associating the individual Hough peaks/Kikuchi bands with their crystallographic index it will be confirmed that bands of consistent intensity across a low

IQ boundary correspond to symmetric twin planes. Crystal planes which are often associated with twinning symmetry for the crystal systems studied in this paper are listed in Table 3-1.

Table 3-1: All planes that display symmetry after twinning and were used for orientation indexing are shown in this table for a number of crystal structures and twins types. Additionally, the number of symmetric planes from the family of planes is shown. The probability represents that number of detected bands in a randomly twinned EBSD pattern that would correlate to twinning symmetry.

	FCC	BCC	HCP Tensile	HCP Compression	HCP Double
Matching plane (# matching/ # possible)	0 1 1 (3/6)	0 1 1 (3/6)	0 1 -1 2 (1/6)	0 1 -1 1 (1/6)	1 1 -2 0 (1/6)
	1 1 1 (1/4)	1 1 2 (3/12)	1 1 -2 0 (1/6)	1 1 -2 0 (1/6)	1 0 -1 3/1-101* (2/12)
	1 1 3 (6/12)		1-101/1-103* (2/12)	1 0 -1 3/1-100* (2/12)	
# of sym. planes	10	6	4	4	3
Total # of planes	27	33	34	34	34
Probability	***37%***	***18%***	***12%***	***12%***	***9%***

3.3 Twin Boundary Measurement Results

This section is divided into three sub-sections based on the three materials that were tested. The first two materials serve as proof of concept for the method as well as providing insights into the method itself. The third sub-section shows how this method is used to identify extremely thin compression twins that are not identified by standard EBSD.

3.3.1 Inconel 600 Twin Boundary Measurements

In scanning the Inconel 600 specimen, high quality EBSD patterns were extracted at several different camera binning levels. The scans contained both regular grain boundaries and also twin boundaries. Many of the twins in the scans were relatively large and easy to detect. Figure 3-4 shows one of the Inconel scans used for this study with easily visible twins. While the twin boundaries can be visually identified based on the extremely straight lines they form, or the

twin plane trace, they can also be detected by the method presented in this study. A plot of individual image quality or Kikuchi band intensity is shown in Fig. 3-5 for a line scan that crosses a twin boundary. In this graph, the intensity of the symmetric planes, namely (1 0 -1) (-1 -3 1) (1 -3 -1) (1 -1 -3), associated with twinning is shown to be fairly constant especially when compared to the other planes not related to twinning. The intensity profiles of the twin related planes experience a slight rise at the grain boundary, but overall the value is fairly constant. This small increase could be attributed to the image processing applied in the collection of the data. Patterns from both sides of the twin boundary are shown in Figure 3-6 along with arrows indicating the matching Kikuchi bands across the boundary.

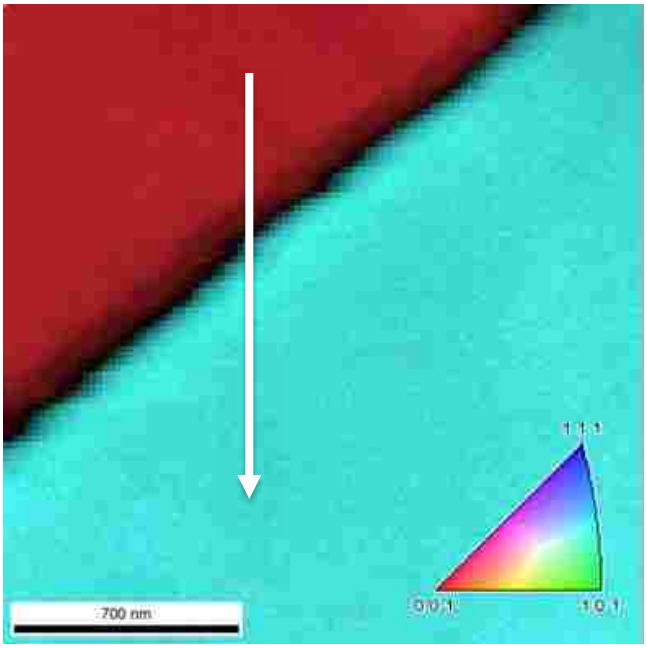


Figure 3-4: Example Inconel 600 twin boundary analyzed. The patterns used to get this data were collected at 2x2 binning.

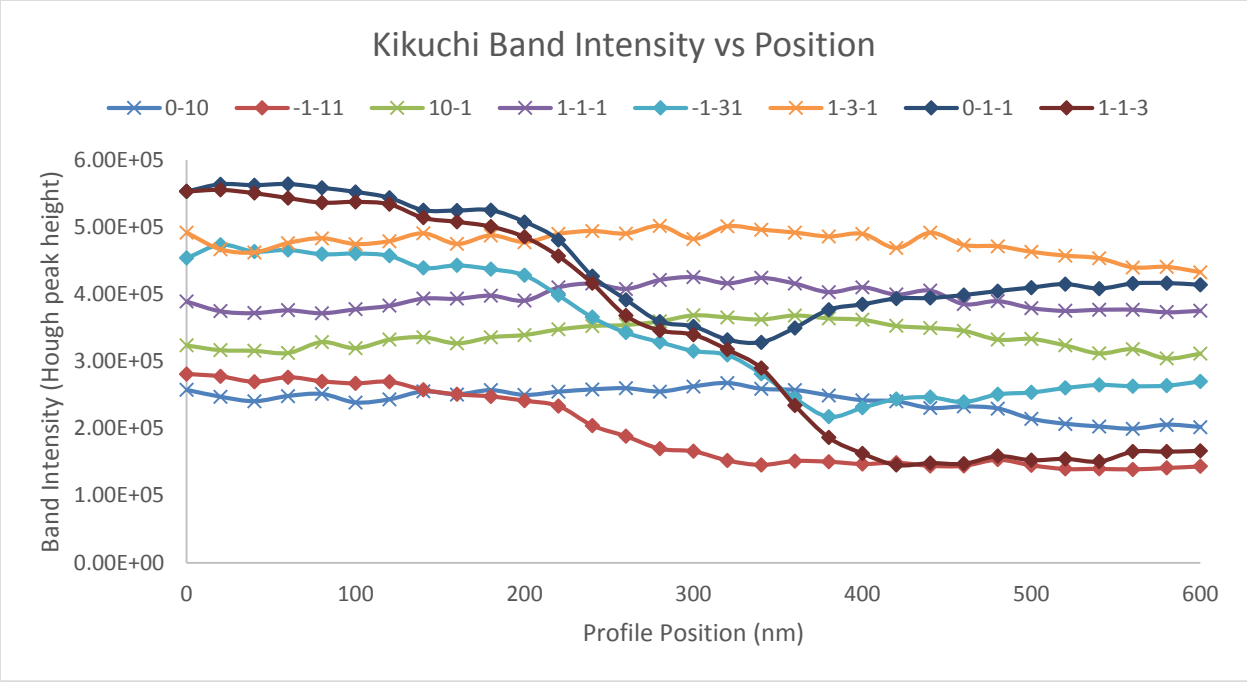


Figure 3-5: Intensity profile of individual Kikuchi bands across an Inconel 600 sample scanned with 2x2 binned patterns. Planes which did not exhibit typical drops in IQ across the boundary have an * in the legend.

.

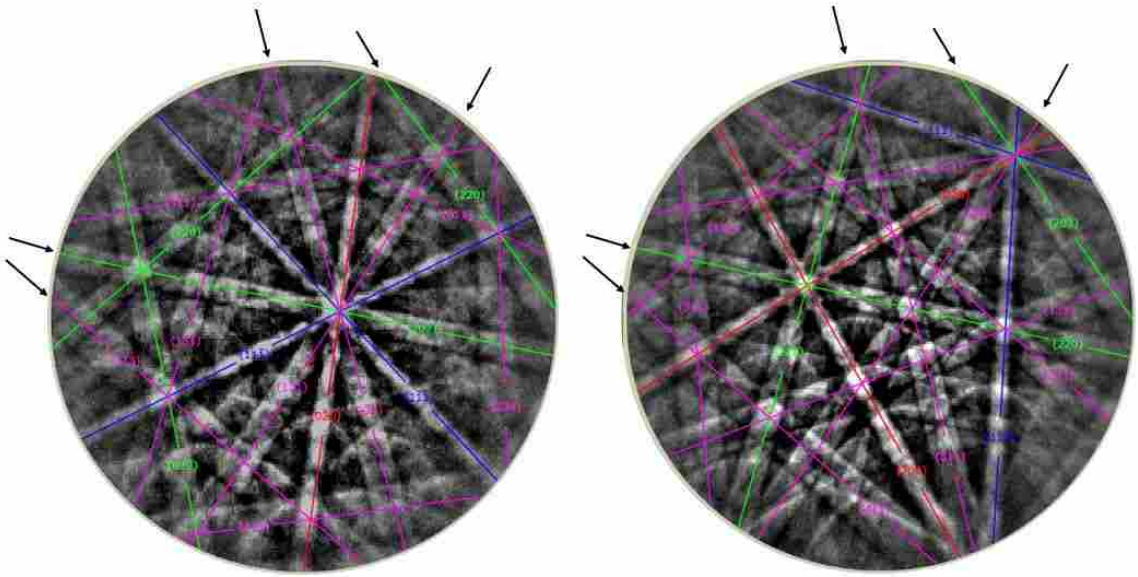


Figure 3-6: Example of Inconel 600 patterns on either side of the twin boundary, parent (left) and twin (right). Indexed planes are overlaid on the patterns. Additional arrows around the outside of each pattern indicate matching Kikuchi bands.

The lack of steady intensity across the twin boundary for the twin plane highlights a critical point in the collection of patterns used for this technique. The patterns for this particular scan were subjected to an image processing routine that normalized the intensity histogram of each pattern in order to increase the visible contrast in the pattern. While this method often makes patterns more visually appealing it can be compared to increasing gain in the detector camera, which increases signal along with noise. The increase in noise can cause especially noticeable issues in grain boundary regions where the signal intensity is at a minimum and the noise is amplified more than in other regions.

In addition to the note on image processing of EBSD patterns for this technique the effect of camera binning was also observed. Patterns of Ni were collected at the following binning levels: 2x2 (240x40 pixels) and 8x8 (60x60 pixels). The results from the 8x8 binning are shown in Fig 3-7 while the 2x2 binning data is shown in Fig 3-5 to illustrate the impact of each binning level on measuring band intensity. As can be seen in the two graph, binning doesn't appear to significantly affect the measurement of individual band intensity. While binning levels beyond 8x8 may weaken the ability of this method to accurately measure band intensity the previous example proves that this technique may be performed on lower resolution EBSD patterns without loss of measurement resolution. This means that the data used for this type of analysis could be collected at extremely high speeds allowing large areas to be scanned or smaller areas with fine step sizes.

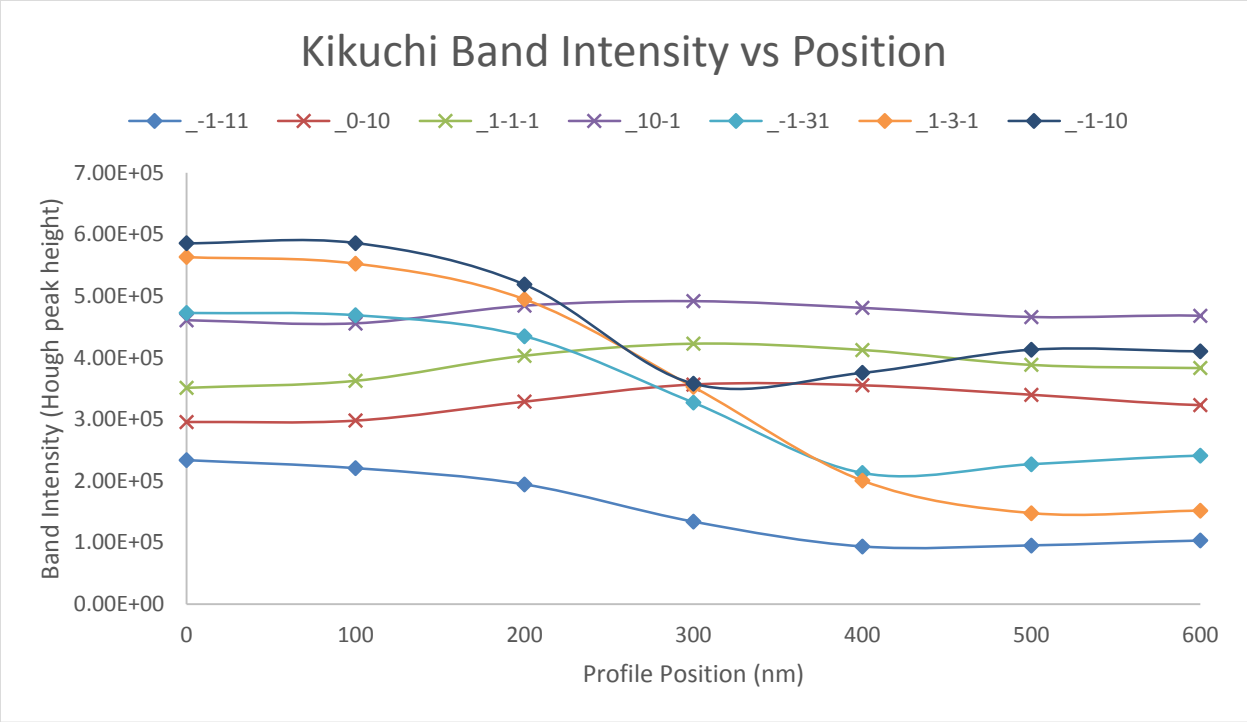


Figure 3-7: Individual Kikuchi band intensity profile across a twin boundary in Inconel 600. This data was collected with patterns binned at 8x8.

For twin boundaries in the Inconel 600 specimen several crystallographic planes, in addition to the twin plane, showed constant Kikuchi band intensity. These planes, which are listed in Table 3-1, include the {011} and {113} planes. Simulations of the equivalent twin rotation in an FCC material, 60 degrees about <111>, reveals that some of these planes are expected to exhibit alignment across the twin boundary. As well as listing the planes that align across a twin boundary, Table 3-1 contains the number of planes pertaining to each family that should align across the twin boundary. In FCC materials 10 of 27 potential bands used for indexing (Only four family of planes were considered for indexing the Inconel 600) would be aligned across a twin boundary. This indicates that for a random twin boundary approximately 37% of the detected bands in patterns crossing the boundary will maintain their intensity. To illustrate the idea of multiple planes aligning across a twin boundary two kinematic simulated

EBSD patterns (parent and twin) were created to show these planes for the Inconel 600 (see Fig 3-8). The example in Figure 3-8 contains 17 Kikuchi bands of which 7 are aligned with each other.

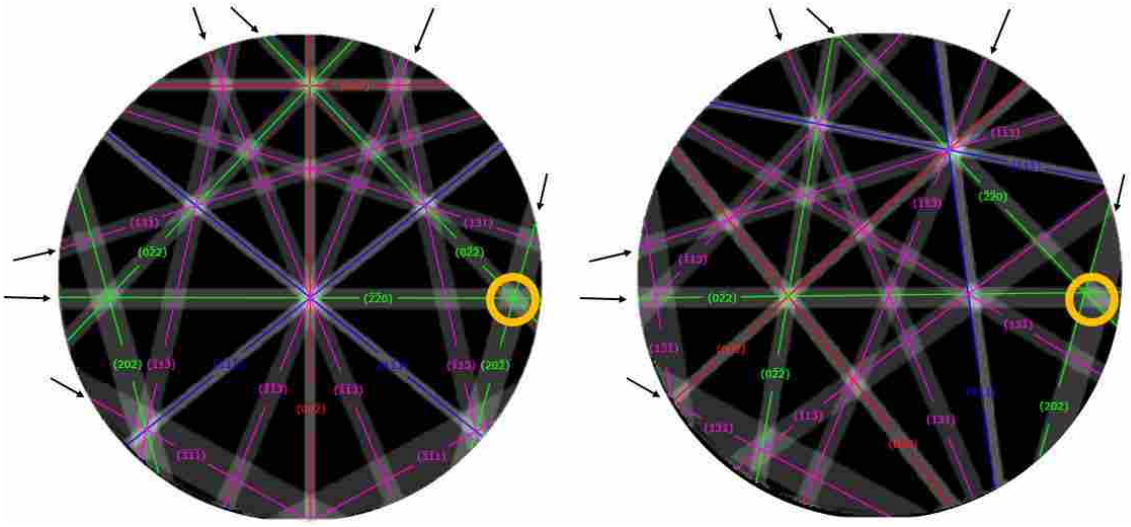


Figure 3-8: Kinematic simulations of EBSD patterns with indexed bands overlaid on both parent pattern (left) and twinned pattern (right). The plane of each Kikuchi band is inline with the drawn band. Matching planes are indicated with arrows around the outsides of the patterns and a matching pole is circled on the right side of both patterns. Note the twin plane is not present in this example.

3.3.2 Tantalum Twin Boundary Measurements

Scans of the twinned Ta specimen also demonstrated consistent band intensity of the twin plane and other symmetry planes ({0 1 1} {1 1 2}), while other planes exhibited the usual plunge in IQ at the grain boundary (Fig. 3-9 shows a subset of the planes). Because the density of Ta is much higher than Inconel 600, the IV is significantly smaller, and smaller step sizes were possible (20 nm). The finer raster size results in better spatial resolution of the Kikuchi band intensity profiles, potentially leading to better detection of smaller microstructural features. It

should be noted that the {123} planes also displayed some form of twinning symmetry, but they were not used to index the patterns, and hence were not included in the band IQ assessment. The planes that maintain relatively constant intensity across the twin boundary are consistent with the plane families listed in Table 3-1. Figure 3-10 shows how some of these planes, in addition to the twin plane, match across the twin boundary.

Results from profiling the individual intensities of several Kikuchi bands confirmed that BCC materials exhibit fewer symmetric twin related planes viewable on the EBSD detector than FCC. Table 3-1 suggests that of the detected bands in a BCC EBSD pattern approximately 18% will be related to symmetric twin planes.

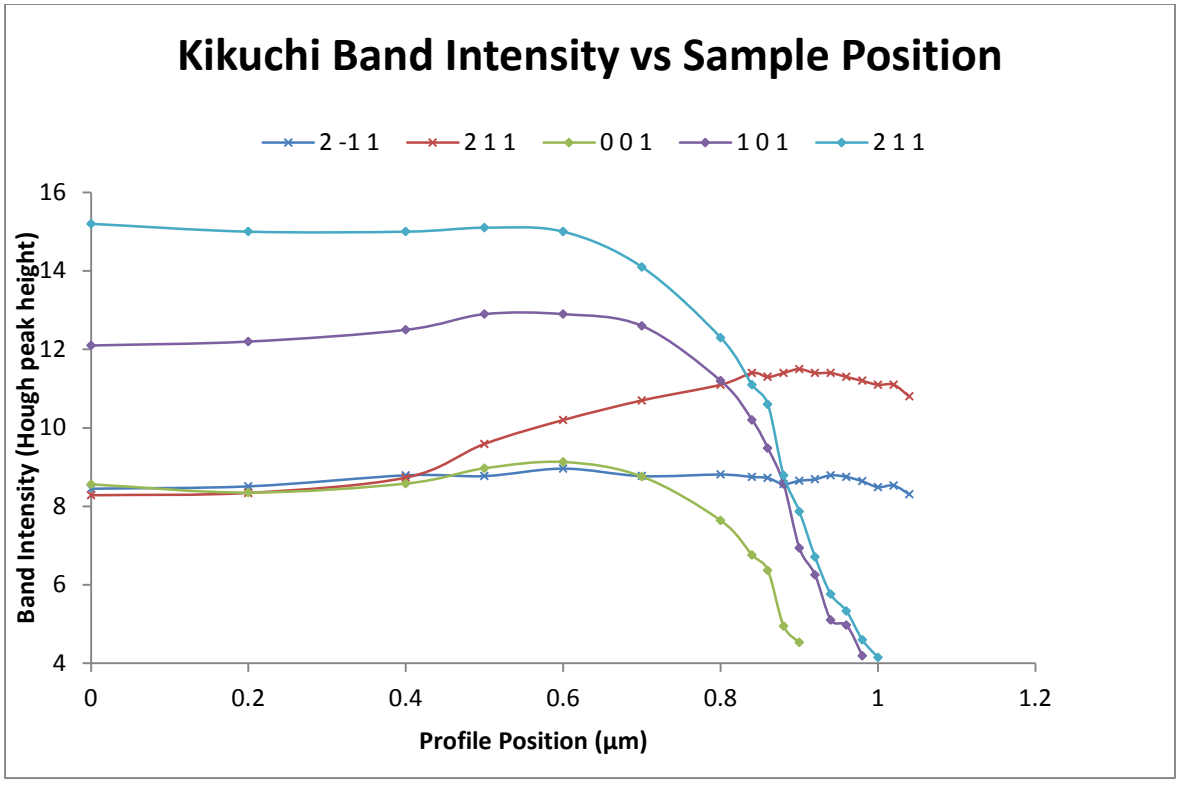


Figure 3-9: Intensity profile of Individual Kikuchi bands for a Ta twin boundary. In this profile only five planes are shown to distinguish the different reactions of individual band image quality near the twin boundary.

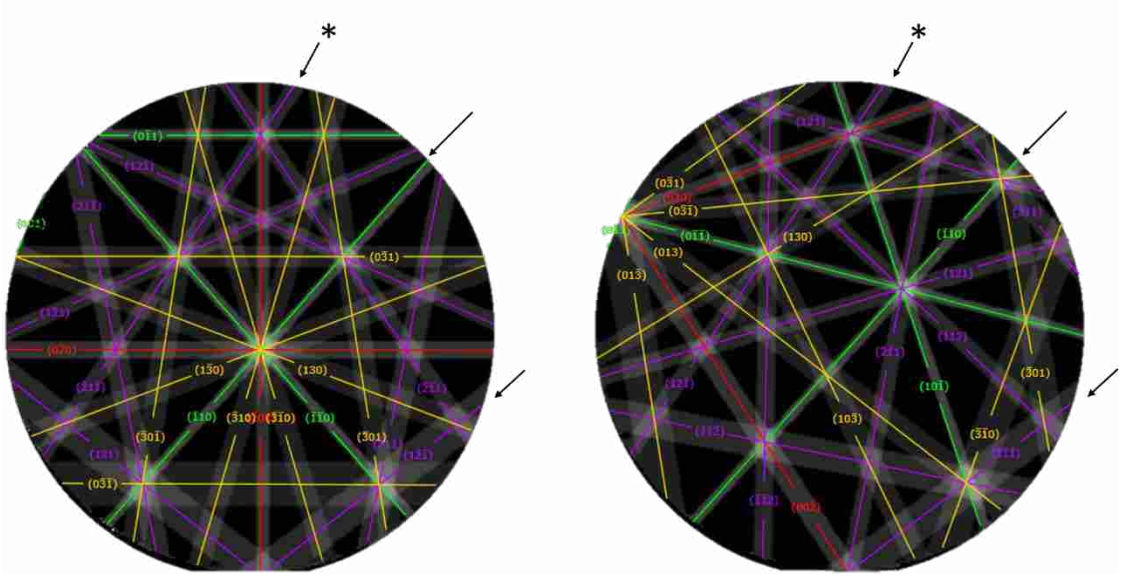


Figure 3-10: Example of a pair of simulated EBSD patterns for ta with a parent pattern (left) and twin (right). The matching planes are indicated with arrows; the arrow pointing at the twin plane is denoted with an *.

3.3.3 Mg AZ31 Twin Boundary Measurements

While scans of the Ta specimens took advantage of the material's high density to achieve high spatial resolution the lower density of Mg inhibits equivalent resolution. The driving force for this study was to provide a method to see the compression twins in AZ31 data collected at a safe step size of 300 nm by correlating them with features in an image quality map. A previous study on forming of this same Mg AZ31 by Scott et al was able to confirm that streaks of lower IQ were generally associated with compression twins (Scott, Miles et al. 2013). That study required the implementation of transmission EBSD which would not be practical for large datasets due to the time and sample preparation constraints.

Although the step size selected was larger than the width of the compression twins in AZ31, sufficient diffraction was still obtained from certain thin compression twins to correctly

identify them with traditional OIM techniques. However, many other compression twins were completely missed by the standard approach.

Validation of the band IQ method was performed on a series of AZ31 compression twins found in Figure 3-11 that were large enough to be properly indexed by traditional OIM. Analysis of the Kikuchi band intensity values across these twin boundaries confirms that intensity levels for twin symmetry planes remained constant. An example of the results can be seen in Figures 3-12 and 3-13. Based upon the success of the method on known twins, the same technique was applied to regions of low overall IQ that appeared to contain compression twins but were not identified by traditional orientation indexing.

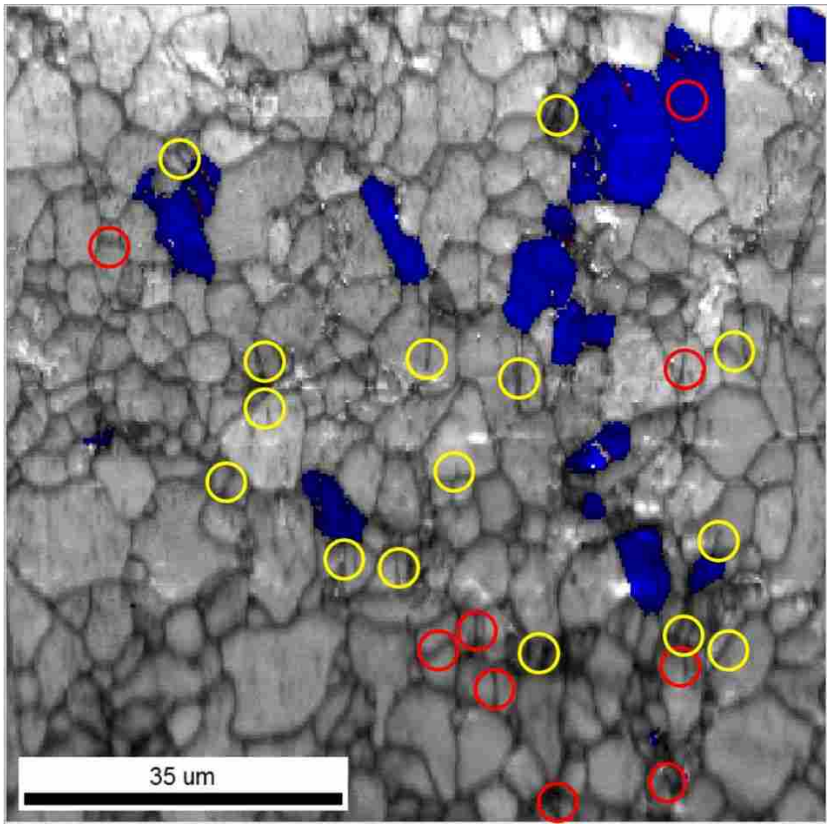


Figure 3-11: Map of IQ along with twinned grains being colored blue and the twins colored red. Circles have been placed around regions that were analyzed for additional twins. Yellow circles show areas that successfully found more twins while red circles failed to identify any twins.

These areas on the IQ map were typically characterized by dark streaks. The earlier study by Scott et al used transmission EBSD to confirm that these streaks of lower IQ generally represented compression twins (Scott, Miles et al. 2013).

Figures 3-14 and 3-15 demonstrate cases where a compression twin was not identified by standard EBSD but was identified by the band IQ technique developed in this paper. Several twins in the microstructure were found using this technique. Some of those areas are identified in Figure 3-11, but not all streaks of lower IQ were positively identified as compression twins. Figure 3-13 shows an example within the same grain of one twin being detected while another remained uncertain.

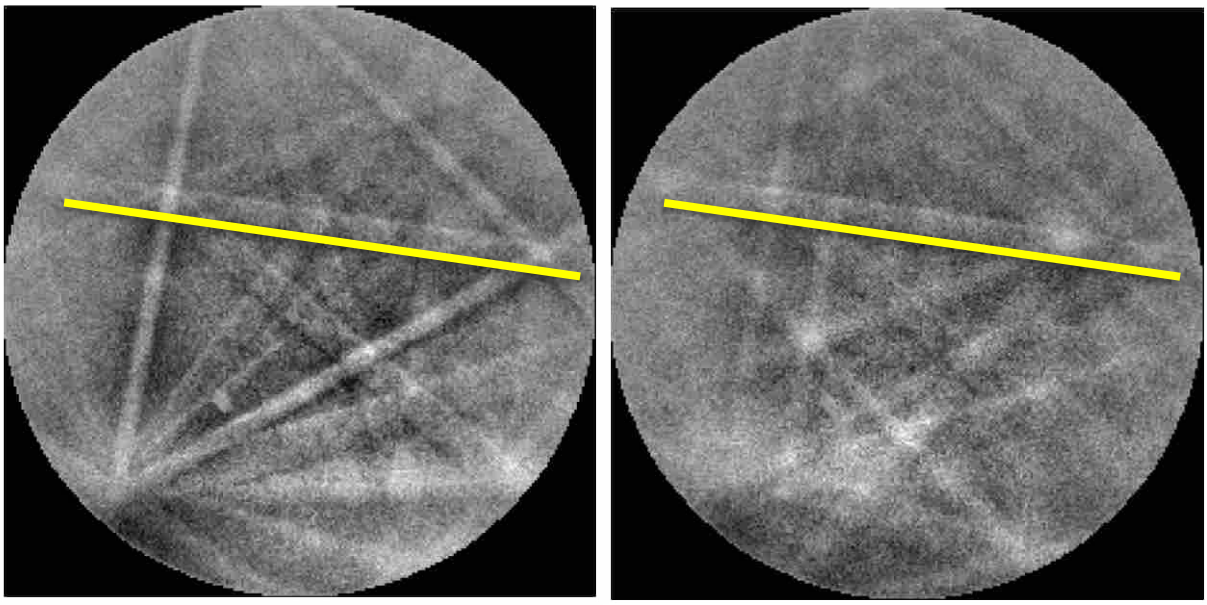
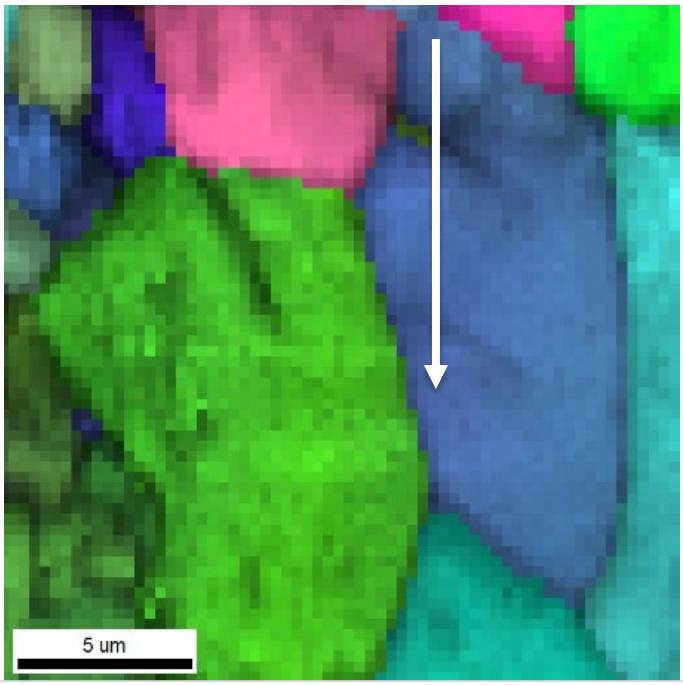


Figure 3-12: Mg AZ31 EBSD patterns from the parent (left) and parent mixed with twin (right). In both patterns one of the symmetric planes which is tied to twinning maintains intensity while other bands experience a noticeable drop in intensity.



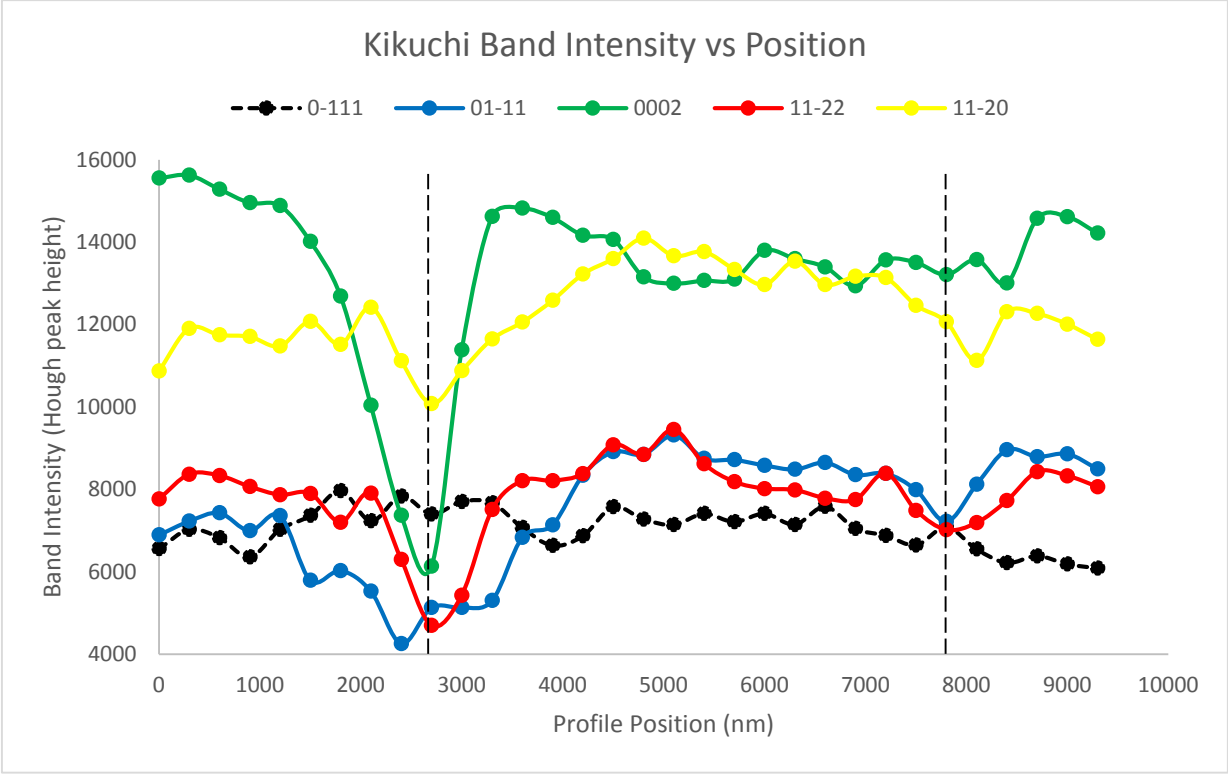


Figure 3-13: Inverse pole figure overlaid with Image Quality map (top) and line profile of individual Kikuchi band intensities (bottom). The profile was taken along the arrow shown in the map. The twin on top was detected by the original data collection whereas the twin on bottom was uncertainly confirmed as a compression twin based on a weak intensity of the symmetric twin-related plane measured.

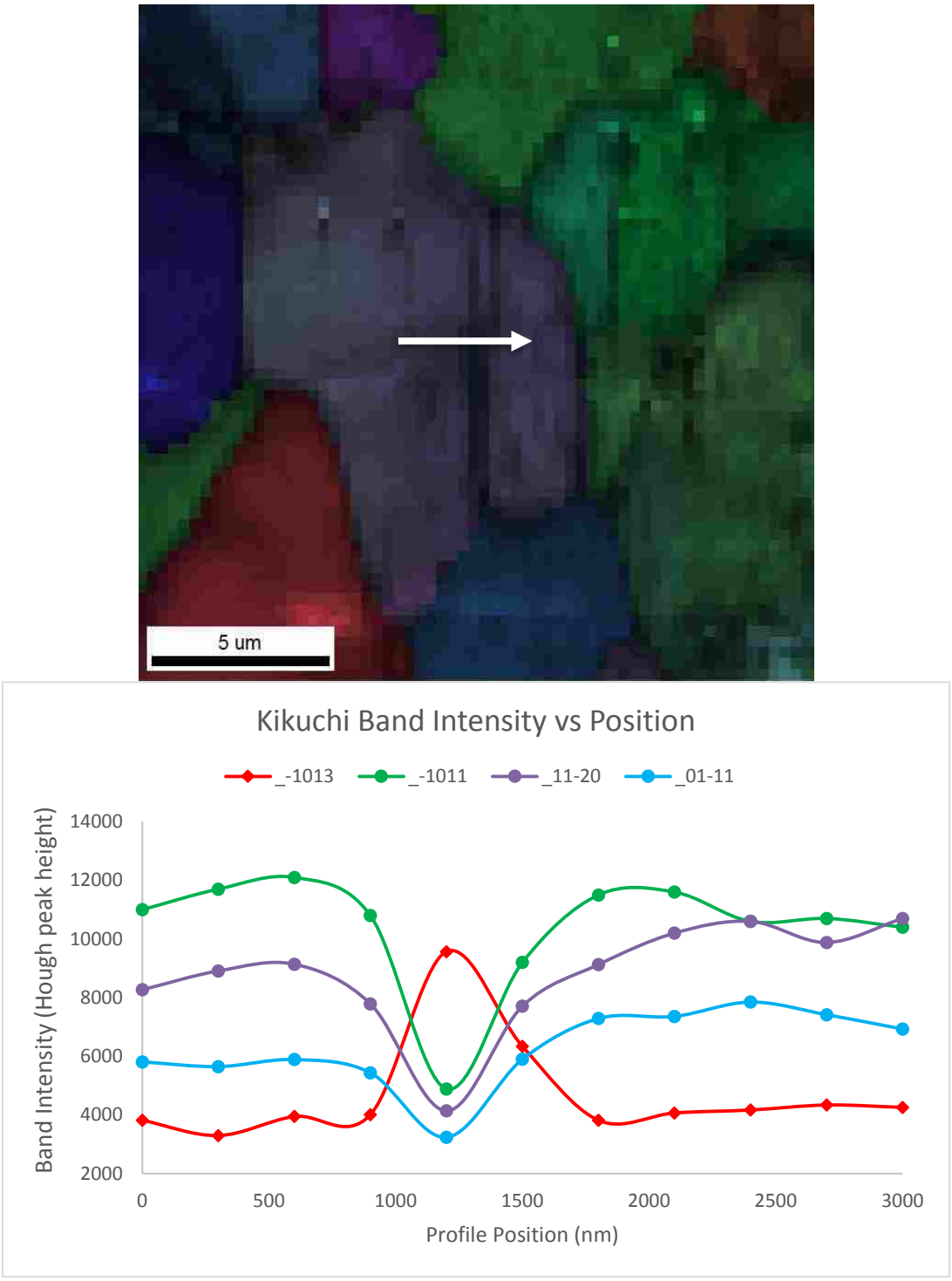


Figure 3-14: IPF and IQ map with initially undetected compression twin (top). Upon observation of the Kikuchi band intensity profile (bottom) one band increases intensity in the region. That band is indicative of twinning at that point due to its special symmetry.

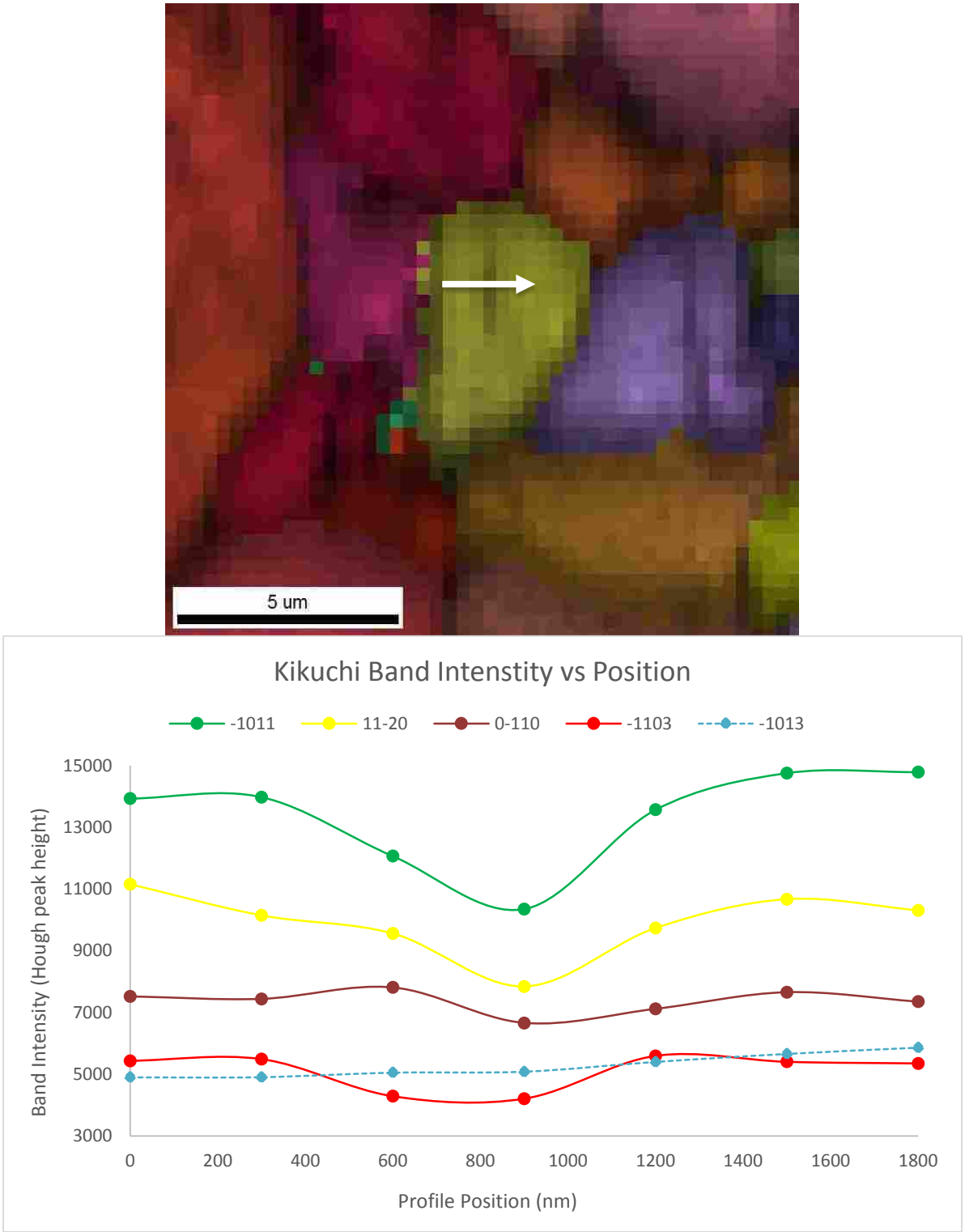


Figure 3-15: IPF and IQ map with initially undetected compression twin (top). The Kikuchi band intensity profile (bottom) has one band that holds its intensity across the low IQ streak.

Of the 24 regions analyzed for the presence of twins 15 were found to contain at least one initially undetected twin. 8 of the other 9 low IQ regions that were checked appeared to correlate mostly with low angle boundaries inside a grain. In the last area which did not detect a twin the low IQ streak aligned with another detected twin in the same grain indicating that the diffraction correlating to twinning was not sufficiently strong to diffuse the parent signal.

More fundamentally, the individual band intensity measure could have failed because the planes required to confirm the presence of compression twins were not aligned to diffract onto the EBSD detector. This possibility exists because unlike FCC materials where 37% of Kikuchi bands in a random EBSD pattern could be related to twinning only 12% of bands in an EBSD pattern of Mg were likely to correspond to symmetric twin planes (see Table 3-1 and Figure 3-15). The potential for missing relevant twin bands on the phosphor was particularly exacerbated by the strong basal texture of the sample (see Figure 3-16).

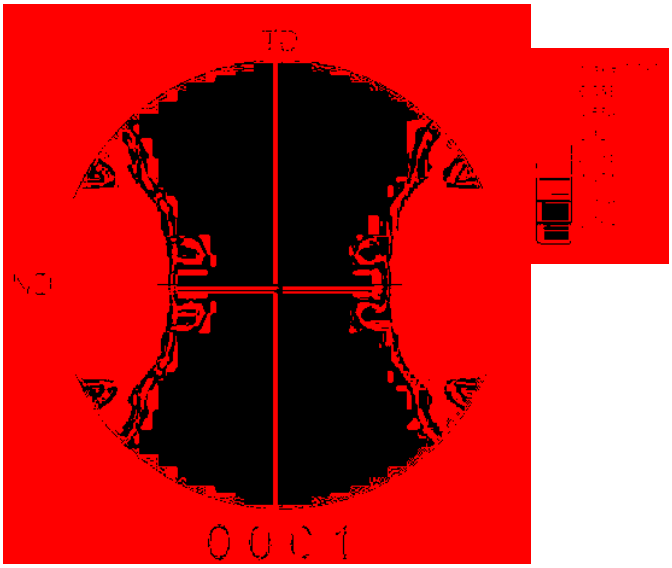


Figure 3-16: Pole figure of AZ31 specimen. Units are times random with a maximum of 9.737 near the ND indicating strong basal texture.

The Mg scan used for this research was taken on the RD (rolling direction) plane of the specimen, such that the c-axis of the grains was fairly well aligned with the horizontal mapping direction. With this sample orientation, some Kikuchi bands, such as the {11-20}, {10-10}, and {10-11} which are critical to identifying twinning, are less likely to appear in the field of view of the EBSD detector. An illustration of this concept is shown in Figure 3-17. The pole figures in Figure 3-17 show the relevant crystallographic planes for twinning. Certain regions of the pole figures are shaded to indicate which bands will not appear on the EBSD detector.

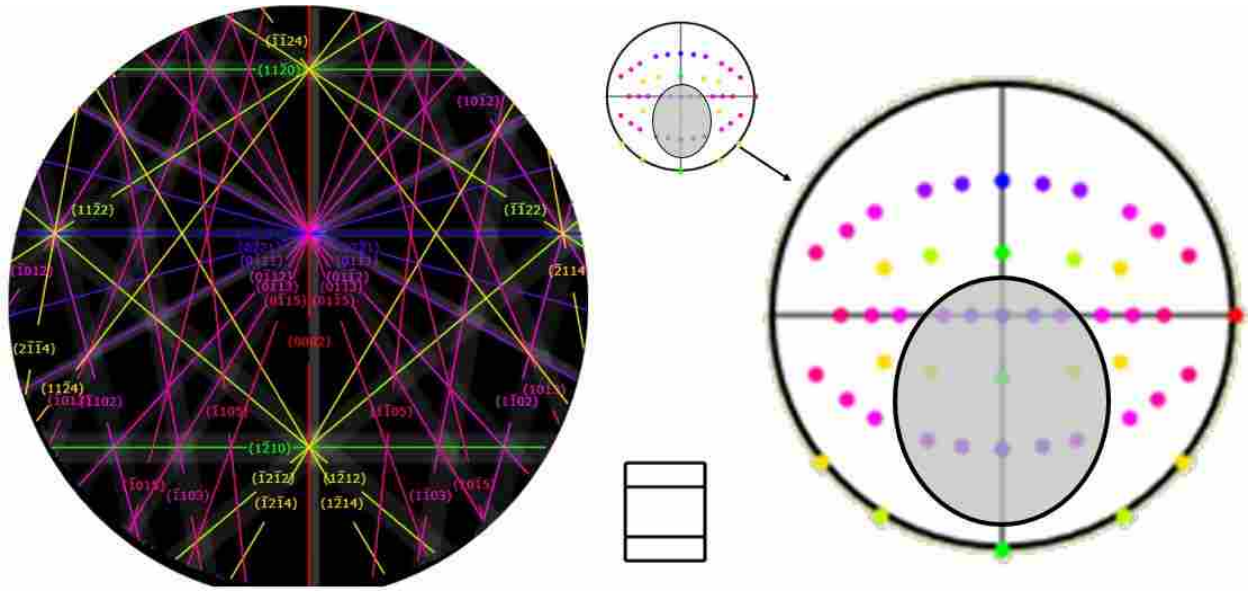


Figure 3-17: Example simulated EBSD pattern of (left) and pole figure (right) of basal textured orientation. Uniquely colored dots on the pole figure represent different families of planes. The shaded gray region on the pole shows which bands would not be seen on the EBSD detector.

The size of the shaded region could be decreased by inserting the detector closer to the specimen; however, inserting the detector closer often results in poor signal near the edge of the EBSD images which would effectively maintain the size of the shaded region. These effects will reduce the already low probability (12%) that symmetric twin planes will be detectable in the

EBSD detector field of view. If the sample were to be rotated 90 degrees for scanning then the likelihood capturing more twin relevant bands would increase as shown by the rotated shaded region. Scanning the ND sample plane would also increase the chance of capturing relevant bands (See Fig. 3-18).

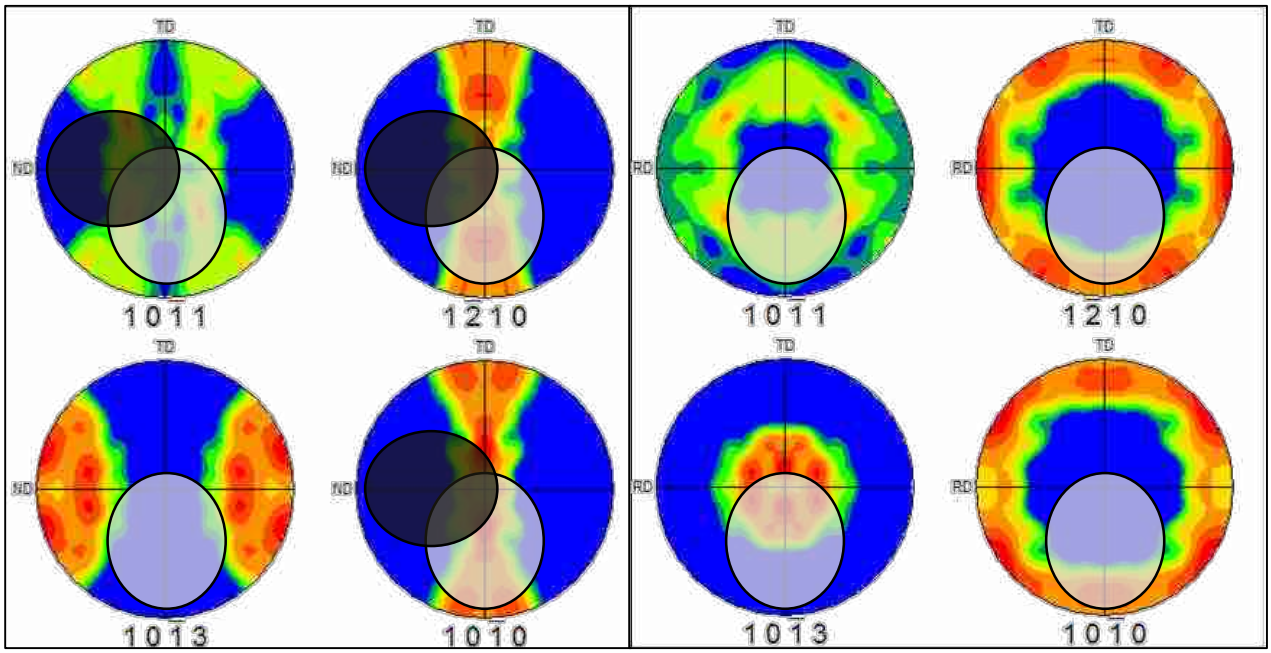


Figure 3-18: Pole figures (PFs) of planes which show some form of alignment across a twin boundary. PFs on the left were measured from the AZ31 specimen as scanned on the RD plane while the PFs on the right were taken by looking at the ND plane. The lightly shaded areas indicate regions where planes cannot be captured by the EBSD detector in the configuration used for this study. Darker shaded regions show the same effect on three PFs if the sample were rotated 90 degrees.

Other factors that could have potentially affected which diffraction bands were present in a basal textured Mg EBSD pattern include deviation of the c-axis from the ND, structure factor of each crystal plane, and the twin variant that is being observed. Some twin variants might be less optimally oriented to diffract aligned planes onto the EBSD detector. The variants considered here are only primary twins; however, secondary twins can also form in Mg, further complicating the identification of twins via this method. The probability of detecting a relevant

Kikuchi band from a secondary twin corresponding to a matching band in the parent grain is only 9%. The only plane seen to align perfectly is the {11-20} plane that pertains to the twin rotation axis; but, the {10-13} set of planes can also line up within two degrees of {10-11} planes as long as the rotation axis of the primary twin remains the same as the secondary twin. Figure 3-19 shows an example of three Mg EBSD patterns (parent, primary twin, and secondary twin) with the one matching plane indicated. In the example, which represents a perfectly basal parent grain, only 1 of 22 Kikuchi band remains constant across all three patterns. This presents a challenge in using band intensities to positively identify secondary twins due to the low probability that the band might be detected. However, as long as one constant Kikuchi band is detected then twins can be said to be present in the region of interest.

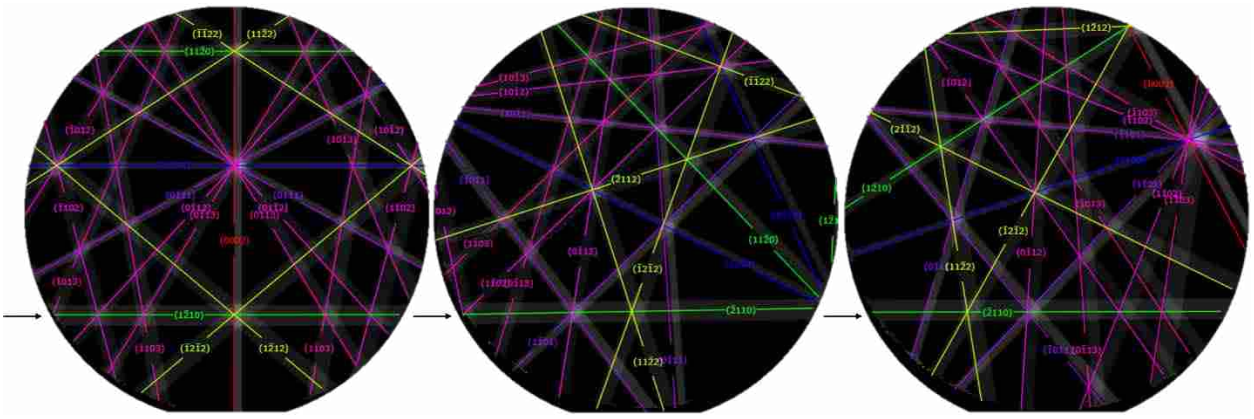


Figure 3-19: Kinematic simulations of Mg EBSD patterns: (left) parent grain, (middle) compression twin, and (right) double twin. In this case of twinning and double twinning only one Kikuchi band matched all three cases. This points to the potential difficulty in using this method for Mg.

3.4 Summary

In this study a technique has been developed which maps the intensity of individual Kikuchi bands in an EBSD pattern across a scan to find twin boundaries in a microstructure.

When the intensity of the Kikuchi bands was tracked across a twin boundary it was found that bands related to planes in the sample with twinning symmetry exhibited a steady intensity across the boundary while other bands experienced a decreased intensity. The method was first applied to Ta (BCC) and Inconel 600 (FCC) to confirm the identification of twin boundaries using this approach. The method was then applied to the identification of thin compression twins in the Mg alloy AZ31, which are often not detectable by standard EBSD techniques due to their small size relative to the electron interaction volume. These twins can often be picked out by observing streaks of low image quality within a grain. When the band IQ method was applied to these regions of interest 15 new twins were positively identified as compression twins in addition to the 18 initially detected by the EBSD software.

In addition to identifying these thin compression twins in Mg a similar approach could be applied to identify other small microstructural features in other materials. Applications could include small grains with CSL boundaries, or other phase boundaries that can be defined by shared crystallographic planes; for example those due to the relationship between austenite and martensite in a deformed stainless steel.

While the previous discussion has focused around identifying particular microstructural features, the technique of tracking band IQ has also gone through proof of concept testing to extract the inclination of a grain boundary. The idea behind this application is to correlate the rate of Kikuchi band intensity decrease to the grain boundary plane inclination. The exact measurement of the inclination would depend on the shape and size of the electron interaction volume requiring parallel work to that presented in (Sorensen, Basinger et al. 2014).

4 INSIGHTS INTO TWINNING IN MG AZ31: A COMBINED EBSD AND MACHINE LEARNING STUDY

4.1 Background

With the increasing demand for materials with high strength to weight ratios in transportation applications, magnesium alloys have emerged as competitive alternatives for structural components (Mordike and Ebert 2001). However, low formability at room temperatures prevents Mg from being a cost effective solution in mass production environments (B.C. Wonziewiz 1967, Klimanek and Pötzsch 2002, N Ono 2003, Yang, Meng et al. 2004, Wagoner RH 2006, Al-Samman and Gottstein 2008, Chapuis and Driver 2010, Chapuis and Driver 2011, Piao, Chung et al. 2012). One issue that leads to this unfavorable forming characteristic is the strong basal texture of rolled Mg sheets that are used in forming. The hexagonal close packed (HCP) crystal structure of Mg, combined with the rolled texture of the sheet material, limits the availability of readily activated slip systems. The absence of easy slip systems to accommodate the applied strain leads to deformation twinning (B.C. Wonziewiz 1967, Koike 2005, Barnett 2007, Meng, Yang et al. 2008, J Wang 2009, Piao, Chung et al. 2012, Barnett 2007). These twins can serve as nucleation sites for cracks (Klimanek and Pötzsch 2002, Meng, Yang et al. 2008, Barnett 2007). The current work aims to achieve an in-depth understanding of twinning phenomena in Mg alloys, in order to establish links between microstructure and formability at room temperature.

The nucleation of a deformation twin within a crystalline material, such as Mg, depends upon so many variables (Meyers, Vohringer et al. 2001, Barnett, Keshavarz et al. 2004, Capolungo, Beyerlein et al. 2009, J Wang 2010, Beyerlein, McCabe et al. 2011, Jonas, Mu et al. 2011) that it may be considered a stochastic event at the level of available information. In order to begin to unravel the complexity behind the physical process, many observations of nucleation events are required. Once sufficient data is available, some sort of data mining process is then necessary to extract correlations linking the available structure and applied field variables to nucleation and propagation events.

Electron backscatter diffraction (EBSD) offers an ideal data collection technique, with the ability to scan hundreds of sample points per second (Adams, Wright et al. 1993, Adams 1997, Maitland and Gholinia 2007). Furthermore, recently developed high-resolution (HR-EBSD) techniques provide new levels of microstructure data that include not only improved quantification of lattice orientation, but also local relative strains and dislocation density that may critically affect twin formation (Wilkinson, Meaden et al. 2005, Villert, Maurice et al. 2009, Gardner, Adams et al. 2010, Basinger, Fullwood et al. 2011).

Machine learning is one approach to mining the large quantities of data that arise from HR-EBSD methods. Already established in various fields as a reliable method for extracting insights and knowledge from vast databases, machine learning provides a framework whereby complex events, such as deformation twinning, can be connected with local structure and field parameters (Yoram and Nahum , Reich and Travitzky 1995, Sha and Edwards 2007, Tompos, Margitfalvi et al. 2007). In this paper, a machine learning framework will be employed to create a predictive model for complex twin events observed by EBSD techniques. Using this approach,

twin nucleation and propagation in Mg alloy AZ31 were correlated with microstructure attributes, and the importance of these attributes to the physical phenomena were investigated.

Based upon an EBSD data source for a decision tree / machine learning framework the goals of this chapter are to:

- Create a model for describing twin nucleation in grains in the Mg alloy AZ31 by extracting attribute based rules,
- Create a model for describing twin propagation across grain boundaries in the Mg alloy AZ31 by extracting attribute based rules,
- Rank the physical factors that most influence twinning phenomena, and
- Setup a reliable machine learning framework for mining EBSD data such that future studies of rare and complex events may be accelerated.

4.2 Materials and Methods

The material studied in this research was a 3 mm thick cold rolled and annealed AZ31 Mg alloy plate with an initially twin-free microstructure. Specimens were first prepared by being cut out of the fully annealed AZ31 plates into 3mm x 4mm x 3mm square prisms using a wire EDM to avoid the introduction of stresses prior to straining. Samples were then prepared for EBSD study using mechanical polishing followed by an OP-S colloidal silica slurry polish. Finally, the sample was etched with a solution of 60% ethanol, 20% distilled water, 15% acetic acid and 5% nitric acid. Using a focused ion beam, platinum fiducial marks were deposited on the specimen surface to measure approximate 2-dimensional loading direction strain in the scan area (Khosravani, Fullwood et al. Submitted 2012).

For this study one specimen was compressed along the rolling direction (RD) up to ~3% strain. This strain was chosen because it was the point at which there was sufficient twin nucleation and propagation for detailed analysis using machine learning, based on prior work (Scott 2012).

EBSD scans were carried out on the FEI-Helios NanoLab^{TM} 600i SEM equipped with OIM^{TM} data acquisition software and a high-speed Hikari^{TM} camera (EDAX 2010). A step size of 300 nm was selected for two scan an areas of 84 x 84 μm^2 and 70 x 270 μm^2. Subsequent post processing of data utilized OIM^{TM} Analysis, Matlab, and HR-EBSD to organize variables for input into machine learning (The Mathworks 2006, EDAX 2010). The extracted variables, their origin, and a brief explanation are listed in Table 4-1.

Certain of these variables were selected to capture morphological data (i.e. grain size, neighboring grain size, GB length, GB orientation, and orientation relative to loading). Such features may affect strain compatibility which is tied to the nucleation of twinning. Variables that fall under the morphology category could be expected to exhibit Hall-Petch type relations since many of them deal with feature size. Feature size plays a large role in both current Hall-Petch type models (Barnett, Keshavarz et al. 2004, Barnett 2007) and stochastic models (Beyerlein and Tome 2010, Beyerlein, McCabe et al. 2011), since larger features present statistically greater areas for events to occur.

Other attributes collected for the machine learning algorithm represent incentive to slip or twin, along with information regarding localization of deformation (i.e. Schmid factors, kernel average misorientation, and local dislocation densities). GND content, measureable by HR-EBSD, indicates stress localization in a material. Stress concentrations are of particular interest when considering twinning because the augmentation of the global stress might be sufficient to

cause twinning to be favored over another form of easy slip in Mg. The kernel average misorientation can also help highlight areas of increased deformation with dislocations and stress concentrations. The Schmid factor of a particular system also serves as an indicator of the likelihood of slip activity.

Table 4-1: Input attributes (parameters) for machine learning. Check marks indicate which attributes were utilized as inputs for creating each model. Note that the highlighted check marks indicate the important microstructural features as found by each machine learning model. (N: nucleation, P: propagation).

Source	***Attribute*** **- Description**	N	P
OIM Analysis	*Grain size* – Equivalent diameter. Calculated as the diameter of a circle with the same area as the measured grain	✓	
	Neighboring grain size – Average neighbor grain size	✓	
	Relative grain size – Grain size divided by neighboring grain size	✓	
	Number of neighbors – number of neighboring grains	✓	
	Deviation of c-axis from RD, TD, and ND – Smallest angle of misorientation	✓	✓
		✓	✓
	Kernel average misorientation – Average misorientation of directly neighboring points (with 5° cutoff)	✓	✓
	Schmid factors (SF) of basal <a> and pyramidal <c+a> slip systems – Maximum value of each slip system taken as a grain average		
Matlab	*Schmid factors of* $\{10\bar{1}2\}\langle 10\bar{1}\,\bar{1}\rangle$ *tensile twins* - Maximum value, taken as a grain average (also considers the possibility of negative values)	✓	
	Ratio of twinning Schmid factor to <c+a> Schmid factor – Maximum SF for twinning divided by max SF for pyramidal <c+a>	✓	✓
	GB misorientation - Rodriguez misorientation		✓
	GB length - length along measured GB segment		✓
	GB straightness – Average distance of the measured boundary from a straight line connecting the boundary endpoints		✓
	Approximate GB orientation - Angle between the GB trace and RD		✓
HR-EBSD	*Local dislocation densities* – Grain average of sum of Nye tensor terms	✓	✓

The remaining attribute of GB misorientation not only reflects morphology (local relative orientations), but also partially reflects energy associated with GBs. This also relates to GB defects, and is likely to affect twin nucleation or propagation (Beyerlein, R.J.McCabe et al.

2011). It has already been observed that twins nucleate at GBs due to the free surface energy that would facilitate twin nucleation. However, not all boundaries have equal energy and thus misorientation might provide a good measure of surface energy which will determine how likely a twin is to nucleate (or propagate) at a particular boundary.

Properties inside the twinned regions were ignored since these areas are not representative of the conditions that resulted in twin activity. Furthermore, grains and grain boundaries near the edge of the scanned area were excluded from consideration. These areas were ignored due to a lack of complete information for various local attributes (such as grain size).

Once the previously described variables were collected for each grain and grain boundary they were organized into data structures suitable for input into a machine learning program. This study used a publicly available machine learning package, WEKA, for the data mining process (WEKA 2010). A J48 decision tree which applies a C4.5 algorithm was then chosen to find correlations between input parameters and twinning (Quinlan 1993). The C4.5 algorithm employs a method to increase the information gain of the data as it is partitioned. Information gain is calculated by the following relations:

$$\boldsymbol{G(S,A) = E(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} E(S_v)} \qquad (4\text{-}1)$$

$$\boldsymbol{E(S) = \sum_{i=1}^{c} -p_i \log_2 p_i} \qquad (4\text{-}2)$$

In these equations S is the total set of instances being considered, S_v is the subset of S with value *v*, A is the attribute under consideration, *v* is a particular value of the considered attribute A, c is the set of possible outputs (2 for this study) for value *v*, and p is the probability

of instance *i* having value *v*. The information gain of selecting a particular attribute is calculated by *Eqn. 4-1* while *Eqn. 4-2* calculates the entropy of moving toward the tested attribute. If the entropy of moving toward one particular attribute is lower than that relating to others, then the information gain is higher for that attribute and the algorithm selects that attribute as the best for the analyzed node. These are the base equations used in the C4.5 algorithm with the J48 tree algorithm making possible continuous as well as discrete inputs.

In the decision tree approach, the root is the first decision in a tree which has a corresponding attribute and cutoff value. The attribute in the root is often considered the most important attribute because it provides the greatest information gain to the model. Branches, similar to the root, are decision points in the tree defined by an attribute with an associated cutoff value. The only difference between the root and branches are position in the tree. Leaves are end points of the tree with the final classifications. In this study the final classifications indicate whether a grain should or should not be expected to nucleate a twin for the first study, and whether a grain boundary will propagate or block a twin for the second study. Relevant roots, branches and leaves are labeled “R’, “B” and “L”, respectively on the decision trees presented in this paper.

Additional constraints were placed on the machine learning algorithm to avoid overly complicated models; these include applying a confidence factor of 0.25 (an algorithm parameter) and disallowing any results with less than two instances (potential noise). In order to create a validated model using a single data set, standard cross-validation techniques were applied to the decision tree algorithm. Cross-validation is a proven machine learning technique whereby the generalization capabilities of the model being made are improved by partitioning the training data into smaller subsets (10 groups per model for this study). The confidence factor value sets a

limit on the minimum acceptable information gain for a branch to be included in the tree. Its values are greater than zero, with larger values requiring less information gain per branch. A minimum of two instances per leaf was also required as an additional precaution to avoid the effects of over fitting (Sha and Edwards 2007, Tompos, Margitfalvi et al. 2007). Over fitting is an occurrence in machine learning that can cause a model to have poor generalization capabilities (i.e. the model only accurately describes a small subset of instances).

The analysis presented in this paper calls for two separate models to be created, one for nucleation of twins in grains and another for propagation of twins across grain boundaries. Nucleation events were characterized by the presence of at least one twin in a grain (i.e. both twins that spontaneously formed in a grain and twins that propagated into the grain from a neighboring grain); and, propagation events were defined by grain boundaries where at least one twin was present on either side of the boundary. The nucleation model took into consideration attributes within a grain, but ignored GB morphology. The propagation model, on the other hand, examined grain boundary events which focused on parameters relating to GBs.

The data collected for this study used detailed inputs from 104 grains to build the nucleation model and 130 grain boundaries from the same specimen (sample compressed along RD to ~3% strain) to build the propagation model. Among the 104 grains used in the creation of the nucleation model, 38 had detectable twins present. The nucleation model took into account a total of 13 attributes including grain size, number of neighboring grains, average size of neighboring grains, relative size of a grain compared to its neighbors, kernel average misorientation, deviation of the c-axis from the RD, TD, and ND, basal Schmid factor, pyramidal <c+a> Schmid factor, tensile twin Schmid factor, the ratio of the twin Schmid factor to the pyramidal <c+a> Schmid factor, and $\log_{10}$ of dislocation density measured by HR-EBSD.

Of the 130 observed grain boundaries that contacted one or more twins, 30 were found to propagate twins. The propagation model took into account a total of 19 attributes including kernel average misorientation on both sides of the grain boundary and the difference (denoted as GB&D), basal Schmid factor (GB&D), the ratio of the twin Schmid factor to the pyramidal <c+a> Schmid factor (GB&D), $\log_{10}$ of dislocation density measured by HR-EBSD (GB&D), deviation of the c-axis from the ND (GB&D), GB straightness, approximate GB orientation, GB length, and GB misorientation.

In addition to the microstructure used to build the nucleation and propagation models another similar microstructure was scanned which underwent 3.5% compression in the RD. This dataset contained 1239 grains and 1127 GBs which were used to check the models. In testing the model created from the smaller dataset on a much larger dataset the generalization ability of the machine learning model will be demonstrated.

In order to quantify the accuracy of the results, each model was analyzed in three distinct ways. First, the overall accuracy of the model was assessed as the total number of correct predictions divided by the total number of instances used to create the model. The target number for this general evaluation was 90% correct predictions; the authors deemed this a suitable number to indicate capturing accurate statistical information while still recognizing the imperfections in each model. Second, the error of each individual leaf was checked by calculating the number of missed predictions in a leaf divided by the total number of instances present in that leaf. This measure was created to help demonstrate which attributes present in the model had the most significance. While no error is preferable any leaf with less than 20% error was still considered a good result. Finally, the individual errors associated with a particular instance within a leaf were checked by visually inspecting maps of microstructure features used

in the machine learning model in order to observe the presence of potentially noisy input data. If reason for noise in the measured data existed then the accuracy of the affected leaf was assumed to be higher than indicated. The combination of these three checks was used to help fully determine the overall strength of the developed models.

4.3 Machine Learning Results

4.3.1 Twin Nucleation in Grains - Results

The model created for nucleation of twins as a function of microstructure attributes is shown in Figure 4-1. In testing the observed microstructure against the created decision tree, 86.5% of the 104 grains and 75.1% of the 1239 grains were correctly categorized by the model.

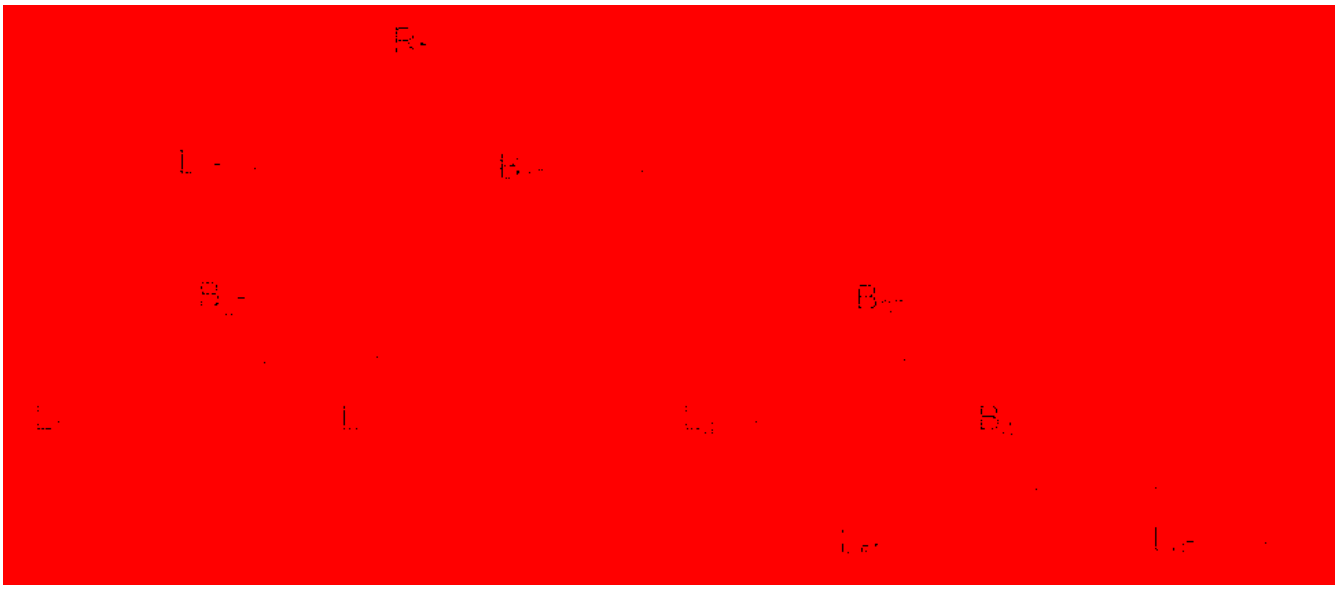


Figure 4-1: Decision tree for characterizing twin nucleation within an individual grain.

This decision tree is relatively shallow, with only five branches (decision points) and six leaves (end points). As seen in Figure 4-1 grain size, dislocation density, c-axis to ND deviation, and basal Schmid factor are the only 4 attributes needed to accurately define twin nucleation within a grain. The presence of grain size in the machine learning model coincides with the findings of many researchers (e.g. (Barnett, Keshavarz et al. 2004)). Furthermore, the location of grain size at the "root" of the tree, and a further appearance at branch B3 within the tree, emphasize its fundamental importance. The next attribute appearing in the decision tree is basal Schmid factor; its appearance is in line with the findings of others, as it indicates whether a grain is "hard" or "soft", and therefore prone to twinning (e.g. (Khosravani, Fullwood et al. Submitted 2012)). The dislocation density and c-axis to ND deviation might be considered new findings pertinent to the twinning phenomenon in Mg. These attributes show up near the ends of the tree, indicating that they don't provide as much information gain as the other attributes, but their contribution is still important.

Additionally, the validation of the model may be visually ascertained by mapping the predictions made in the decision tree to the grain structure tested (see Figure 4-2). This map was then checked against other structure/property maps (Figure 4-3), as described earlier, to draw conclusions about the ability of the model to predict correctly. Such an analysis helps to highlight areas of potential information noise and how well they were dealt with by the machine learning algorithm.

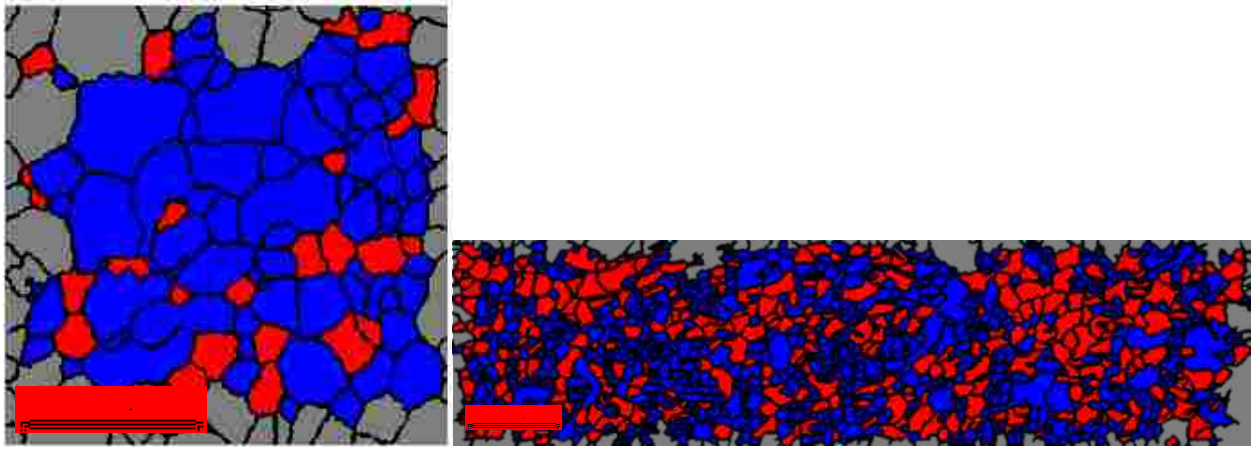


Figure 4-2: Error maps of decision tree for predicting twinning in individual grains. Microstructure used to build model, 104 grains (left) and test generalization, 1239 grains (right). Correct predictions (blue) and incorrect predictions (red) are shown except for edge grains (gray) which were excluded due to incomplete information.

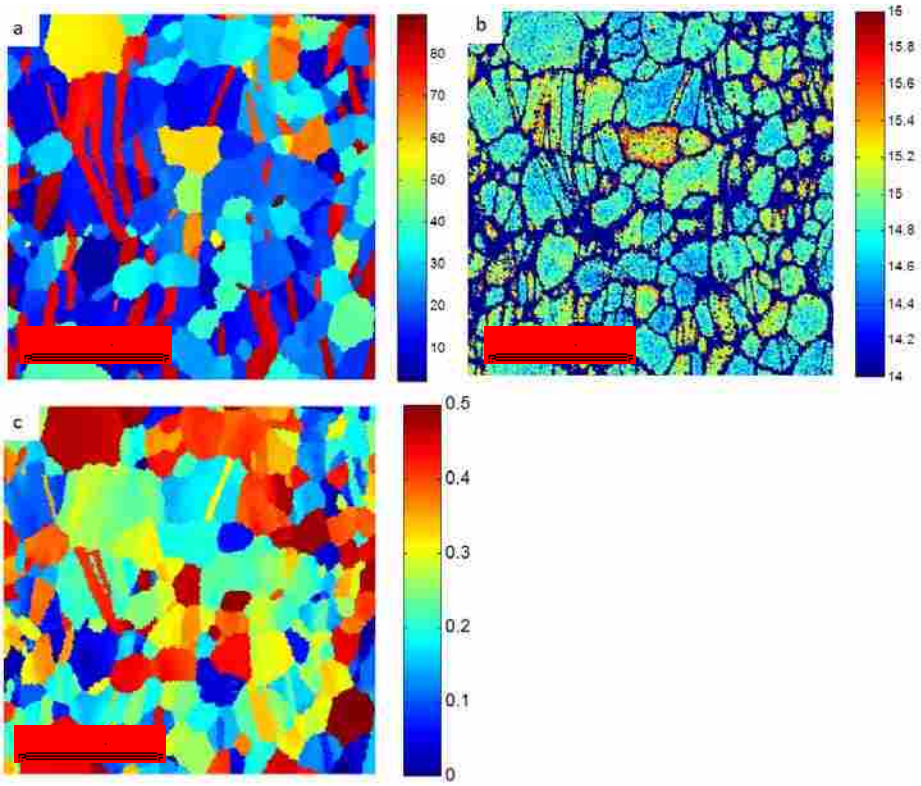


Figure 4-3: Feature maps. ND deviation (a), $\log_{10}$ of dislocation density (b), and basal Schmid factor (c). Areas of dark blue in the dislocation density map indicate noisy, neglected points.

4.3.2 Twin Propagation across Grain Boundaries - Results

The decision tree model for twin propagation is shown in Figure 4-4. Testing for twin propagation across grain boundaries resulted in an accuracy of 96.1% for the 130 twinned GBs and 75.5% for the 1127 GBS used for testing generalization. However, the higher accuracy came with an increased depth of decision tree. Whereas the nucleation of twins within grains had a relatively small decision tree (5 branches and 6 leaves) the model for twin propagation had 11 branches and 12 leaves, approximately twice the size of the prior tree.

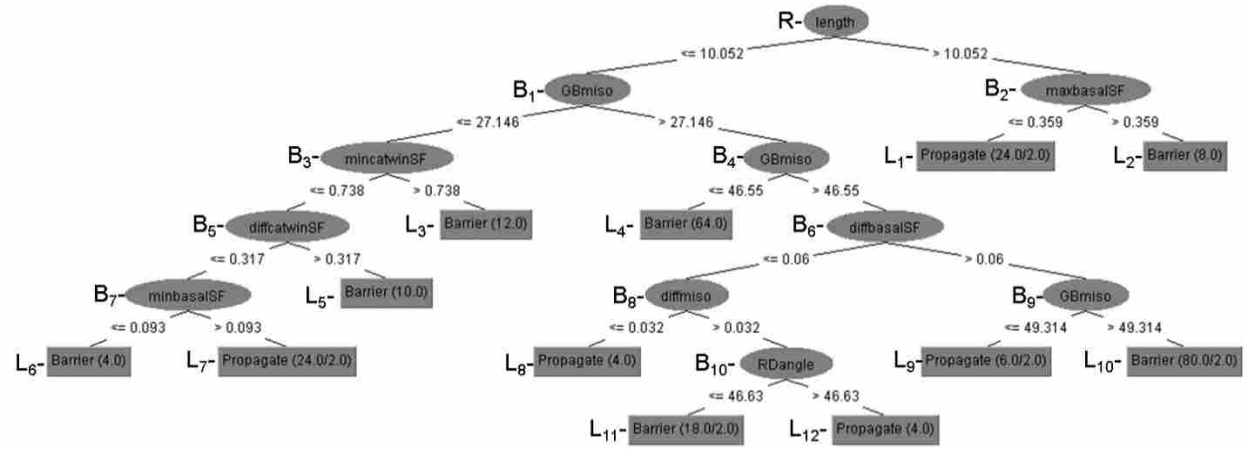


Figure 4-4: Decision tree for characterizing twin propagation across grain boundaries.

Within the propagation decision tree 9 of the possible 19 distinct attributes are present: GB length, GB misorientation, minimum twin to <c+a> Schmid factor ratio, difference of twin to <c+a> Schmid factor across the GB, minimum basal Schmid factor, difference of basal Schmid factor across the GB, difference of kernel average misorientations across the GB, the approximate angle of the GB on the sample surface relative to the RD, and the maximum basal Schmid factor. These 9 attributes stem from 6 basic properties that are highlighted in Table 4-1.

In that regard, the GB propagation tree is not significantly more complex than the grain nucleation tree.

Several of the attributes in the model reflect findings from previous work. For example, GB misorientation was considered by both Beyerlein et al (Beyerlein, R.J.McCabe et al. 2011) and Khosravani et al (Khosravani, Fullwood et al. Submitted 2012). The basal Schmid factor shows up several times in the tree, confirming the finding by Khosravani et al. that showed the importance of basal activity in determining twinning criteria. Other attributes such as GB length captured multiple aspects of prior research such as the grain size effect observed by Barnett (Barnett, Keshavarz et al. 2004) and the statistical nature of twinning at a boundary observed by Beyerlein, since larger boundaries have more potential twin sites.

In order to visualize the accuracy of the GB propagation model, a different type of map is required from that shown in Figure 4-2 (see Figure 4-5). In this case, maps of the six base properties were used, as opposed to all 9 present in the decision tree, to condense the number of examined plots where possible (see Figure 4-6 and Figure 4-3). For instance, GB length did not require its own map since the boundaries can be easily seen in all of the other maps. Additionally, GB misorientation could be partially inferred from the inverse pole figure (IPF - Figure 4-6b) map by observing contrast between grains.

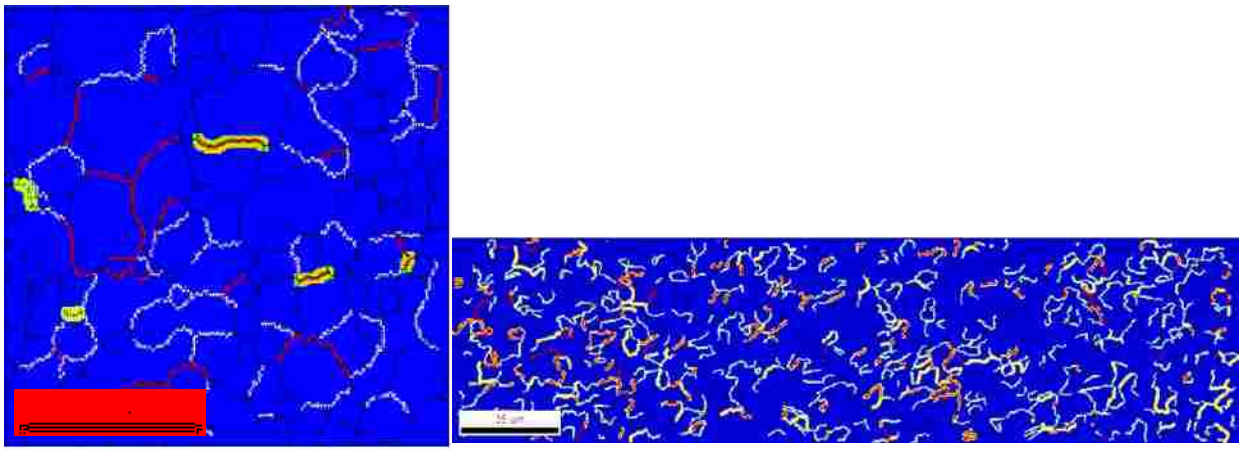


Figure 4-5: Error maps of decision tree for predicting twin propagation across grain boundaries. Microstructure used to build model, 130 GBs (left) and test generalization, 1127 GBs (right). GBs that were predicted to allow propagation of twins are shown in red while those that predicted barriers to twin propagation are pale. Predictions are highlighted in orange for incorrect propagation assignments and yellow for incorrect barrier assignments. Grain boundaries are displayed in black for visualizing the morphology in the model microstructure.

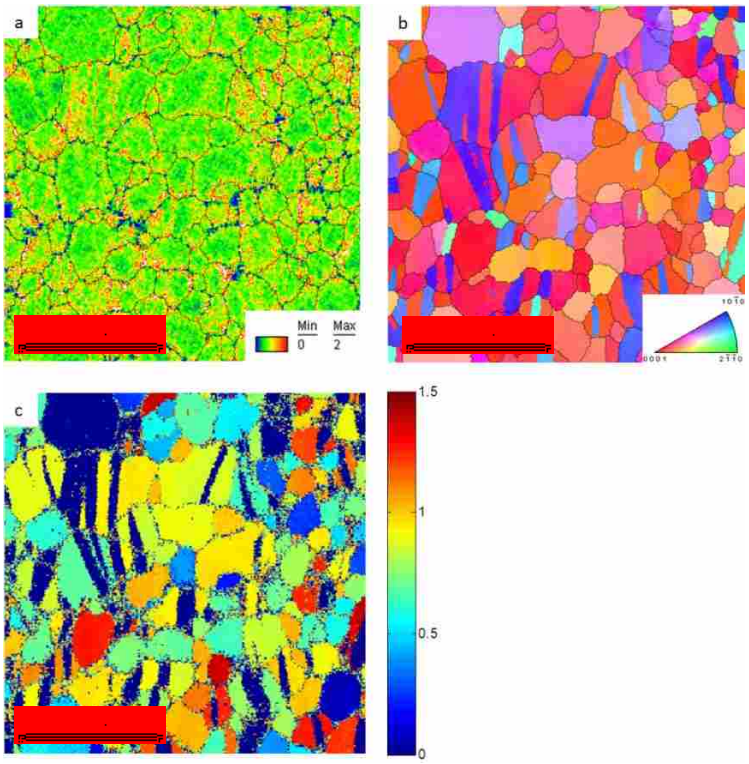


Figure 4-6: Feature maps of additional attributes used in the twin propagation model. Kernel average misorientation (a), inverse pole figure (b), and twin to <c+a> Schmid factor (c).

4.4 Discussion of Machine Learning Model and Results

4.4.1 Twin Nucleation in Grains – Analysis of Model

In analyzing the results and accuracy of the decision tree models, the potential benefits include more than just an accurate twin prediction model. As previously described, these results can serve as a framework for gaining additional insights into the causes of twin nucleation and propagation. An analysis of the grain nucleation model (Figure 4-1) reveals physical phenomena highlighted by each level of the tree. Figure 4-7 will be used to help explain some of these phenomena, but it should be noted that the charts serve to show some functional dependence which is not captured in the decision tree.

At the root (R) of the decision tree is grain size which splits the data path into a leaf (L_1) and a branch (B_1). Leaf L_1 serves to illustrate the difficulty of twin nucleation in relatively smaller grains (≤7.2 μm), consistent with the Hall-Petch approach on the twinning shear stress described by Barnett. Additionally, the machine learning model aligns with the stochastic view of twinning where smaller grains would be less likely to nucleate twins due to fewer available nucleation sites. This relationship can be seen in Figure 4-7c where increasing grain size results in increased likelihood of twinning. Leaf L_1 has a relatively high error (12 of 65 predictions were incorrect), which highlights one negative aspect of the decision tree approach. Clearly from Figure 4-7c, there is a non-linear relationship between grain size and twin occurrence. However, the decision tree creates a discrete division based on grain size to segregate grains that are more likely to twin from those that are less likely to twin. More sophisticated machine learning algorithms can be used to better capture the non-linear functional dependence (for example, neural networks which is skill-based); however, such algorithms are not as good at quantifying

the relative importance of the attributes, and hence at clarifying the physics underlying the phenomena of interest.

The B_1 branch in Figure 4-1 partitions the remaining instances into two sets where the basal Schmid factor is greater than ~0.29 or less than ~0.29. This suggests that larger grains (segregated from the smaller grains at the root of the tree) tend to twin when the maximum basal Schmid factor is less than this cutoff value, whereas those with maximum Schmid factors greater than ~0.29 tend not to nucleate twins. Once again, the relevant chart in Figure 4-7a displays a more complex relationship between Schmid factor and twin probability; nevertheless, the simple cutoff value approach in the model demonstrates that lower Schmid values correlate with a higher tendency to twin, owing to the lack of appropriately oriented basal slip systems to accommodate strain.

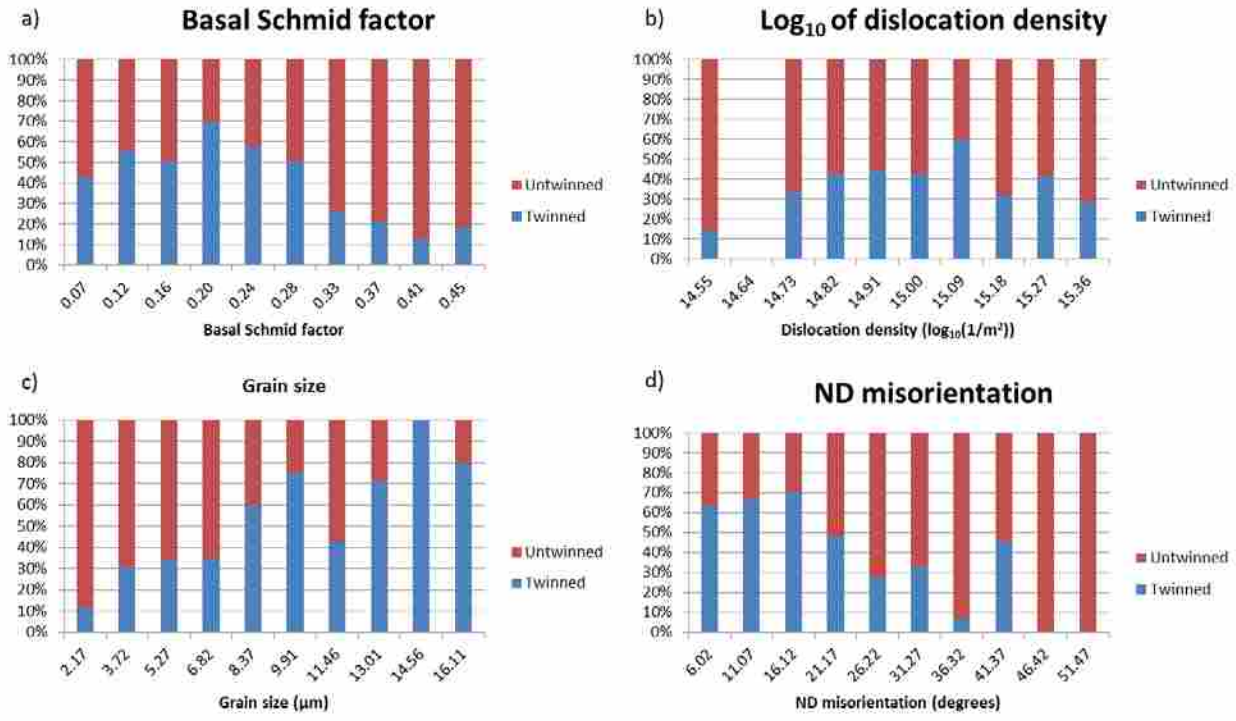


Figure 4-7: Bar charts of relevant features used in the decision tree for the twin nucleation model: basal Schmid factor (a), dislocation density (b), grain size (c), and ND deviation from the c-axis (d).

While grain size and basal Schmid factor appear to be the most influential factors in twin nucleation, a more complete model, as indicated by the decision tree, would require the added consideration of dislocation density and deviation of the c-axis from the sample normal (ND). The dislocation density dependence is apparent at branch B_2, immediately below the basal Schmid factor branch (B_1). When the dislocation content is high enough (2.23e15 m^{-2}) grains tend not to twin, which is likely the result of greater slip activity reducing the necessity for strain accommodation by twinning. Such dependence may not have been immediately apparent from a simple view of the relationship between dislocation density and twin activity, as seen in Figure 4-7b. The strong correlation comes to light only after the small grains, and those with low basal Schmid factors, are filtered out using the decision tree approach.

The additional consideration of c-axis to sample-ND deviation contributes the last unique attribute found in the machine learning model, and is found in branch B_4. This attribute is used as a measure of how strongly aligned a grain is with the basal texture found in the tested samples (i.e. distinguishing "hard" and "soft" grains). In the decision tree this final attribute is applied to a specific set of grains: grains larger than 9.66 μm and with max basal Schmid factors greater than ~0.29. The final leaves (L_5, L_6) of the model demonstrate that grains which are closer to the basal texture tend to twin more, presumably because of the need to accommodate c-axis expansion by means other than <a> type slip. This observation aligns well with the distribution of twinned grains leaning towards "harder" or more basal grains in the relevant subplot (d) of Figure 4-7.

While the model created by the machine learning process reveals important qualitative and quantitative results, there are limits that must be recognized. The inaccuracies in the model might be categorized into two types: first, poor fit of the chosen algorithm to the supplied test

data, and second, noise/inaccuracies in the data used to create the decision tree. The former will be discussed later in the paper along with other effects of using machine learning. The issue of noisy data, however, brings to light the importance of high quality data collection. While the methods associated with EBSD are reliable, they are for this study, only two-dimensional data sets. Such an analysis might cause errors in calculating certain attributes. On average, stereological values of grain size have been shown to be correct (Wright 2010), but in this study, some grain surface areas in the 2D sample section are not accurate enough to make a correct division in the decision tree. While this issue was considered negligible when the above machine learning analysis was performed, it might very well be the cause of complications in accurately predicting twin nucleation in small grains (which may simply be the cross section of the apex of a large grain in the 2D sample). However, it could also be that the nucleation of twins in relatively small grains ($\leq$7.2 μm) is more complex than the machine learning algorithm could accurately predict given the size of the data set or the attributes considered in the model. Regardless, the decision tree created for twin nucleation in a grain was highly successful and served to provide an important set of insights into the phenomenon.

4.4.2 Twin Propagation across Grain Boundaries - Analysis of Model

While the model presented for finding grains that nucleate twins was relatively simple, the decision tree for characterizing propagation of twins across GBs (Figure 4-4) involved several more parameters and was more complex. This is also highlighted by the graphs of underlying dependencies for various attributes, as shown in Figure 4-8. Therefore, the focus of the analysis of this model will not be to explain the tree in its entirety but to extract as many reliable conclusions as possible.

The first attribute that appears at the root of the decision tree (R) for propagation is GB length which was not directly considered by most previous researchers. However, boundary length (or area) should be examined since larger areas statistically increase the probability of twins nucleating or propagating across a boundary. Furthermore, longer grain boundaries seem to propagate twins more readily than their shorter counterparts (≤10.1 μm), indicating a Hall-Petch type effect, or a stochastic-based model, both previously discussed earlier.

Next in the decision tree is GB misorientation (B_1), which had been previously studied by other researchers (e.g. (Beyerlein, R.J.McCabe et al. 2011)). In the machine learning model, GB misorientation appears at three distinct branches (B_1, B_4, B_9) and only applies to GBs shorter than ~10.1 μm, pointing to a more complex relationship between misorientation and twin activity than might previously have been captured by current models in the literature. In general, grain boundaries with lower misorientation are more likely to propagate twins while higher misorientation grain boundaries serve as barriers to twin propagation. The machine learning model identified 27.1^0 as the approximate threshold angle of misorientation below which propagation is favored, showing good agreement with the 25-39^o found by Khosravani et al (Khosravani, Fullwood et al. Submitted 2012). Note that this apparent correlation may also be due to the fact that hard grains neighboring soft grains tend to have high angle boundaries, and that, strain accommodated in the soft grain by slip must be accommodated in the hard grain by twins (see the discussion in (Khosravani, Fullwood et al. Submitted 2012)). However, the overall effect of GB misorientation on twin propagation is fairly complex, as seen in Figure 4-8b. This complexity is expressed in the decision tree model by the reoccurrence of GB misorientation.

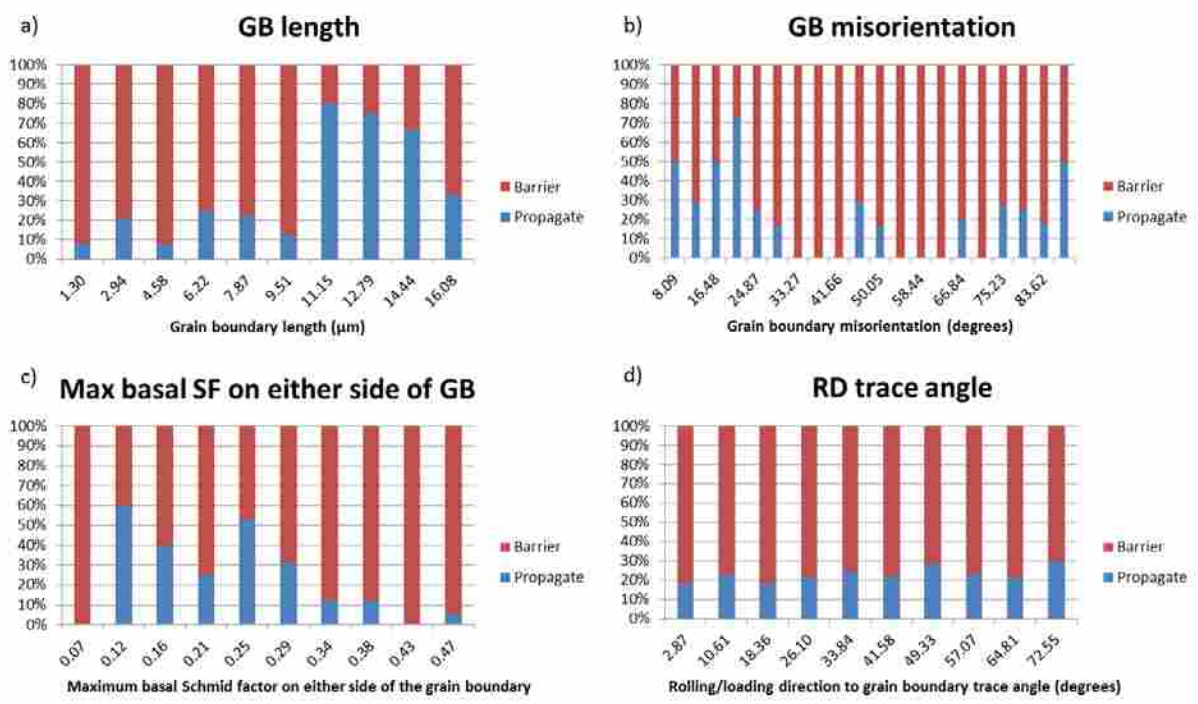


Figure 4-8: Bar charts of relevant features used in the decision tree for the twin propagation model: grain boundary length (a), grain boundary misorientation (b), maximum basal Schmid factor (c), and the angle between the grain boundary trace and loading direction, RD (d).

Another significant attribute in the machine learning model is the basal Schmid factor near grain boundaries (B_2, B_6, B_7). As presented in the decision tree, the maximum basal Schmid factor is the deciding factor in whether longer grain boundaries block or propagate twins. This can also be visualized to some extent via the basal SF plot in Figure 4-8c. According to B_2 when the maximum basal SF on either side of a boundary is less than ~0.36 twins propagate. This leads to the conclusion that if basal slip systems cannot easily accommodate the stress state caused by a twin on the opposing side of a GB then twinning occurs in order to accommodate that stress. The reasoning behind this may be tied to the stress concentration caused when a twin intersects a GB. By assuming constant strain across twin and parent the resulting stress state in the twin will be different from that in the parent matrix. This would cause a sudden change in the

stress state along a grain boundary which it intersects (Figure 4-9), much like the stress concentration at a triple junction. Previous research has demonstrated the stress rise at GBs and triple junctions using finite element modeling for various materials and triple junction configurations (Yue 2003, Fallahi and Ataee 2010). In these studies the stress at a triple junction has been shown to be as much as 30% higher than the average stress in the grain. Referring back to Table 1-1 this increase in stress indicates enough change to allow a transition from activation of prismatic slip (8-10 MPa) to activation of tensile twinning (11-12 MPa). Furthermore, Koike (Koike 2005) discussed the formation of twins due to stress concentration and Barnett (Barnett 2007) showed that twin intersecting other boundaries can lead to void/crack nucleation which also serves to illustrate the increased stresses at these locations. In brief, the stress concentration caused by a twin intersecting a GB is sufficient to increase the local stress such that the stress required for tensile twinning is achieved causing that twin to propagate across the boundary.

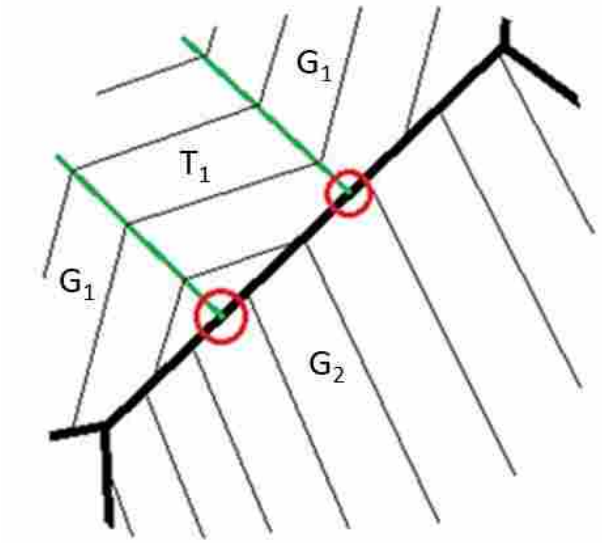


Figure 4-9: Schematic of morphology leading to stress concentration caused by twin intersection at a grain boundary. The areas of stress concentration are circled.

Other attributes present in the decision tree include the ratio of twin Schmid factor to <c+a> Schmid factor (B_3, B_5), kernel average misorientation (B_8), and the average angle between the GB and loading direction (B_{10}). The ratio of Schmid factors (B_3, B_5) may give information as to whether incompatibility will exist at a GB which would then be accommodated by a twin. For example, smaller differences between these ratios across a GB may encourage twins to propagate, since both grains are more likely to twin than activate <c+a> slip. Additionally, kernel average misorientation (B_8) has been used by some researchers as an estimate of dislocation activity (Marion Calcagnatto 2010). With this attribute, smaller differences across a GB tend to result in propagation of twins, and larger differences serve as barriers. Similar observations were made by Khosravani, in that dislocation pileups can correlate with poor twin propagation as well as spontaneous twin nucleation (Khosravani, Fullwood et al. Submitted 2012). The final morphological measure of grain boundary trace angle (B_{10}) seems to show that boundaries more perpendicular to the RD will propagate twins, likely due to the increased area across which the incoherent twin boundary spans. This relationship cannot be seen in Figure 4-8d, but is extracted through the machine learning algorithm pointing to the advantage of the method presented in this paper.

The inaccuracies associated with the model of twin propagation across grain boundaries are small and probably due in part to the large size of the decision tree. Large trees / models tend to increase the likelihood of over fitting which limits the models ability to accurately characterize all data sets. So while the model predicts high accuracy it might be an artifact that could only be clarified by analysis of a much larger data set. Regardless of the potential errors in this model, attributes nearest to the root of the tree would most likely remain the most important, thus maintaining the relative accuracy and interpretation of the model.

4.4.3 Machine Learning Framework - Discussion

When selecting the decision tree algorithm it was understood that it would provide a rule-based description of the studied events related to twinning in AZ31, and not a knowledge-based description. Such a result may be considered a weakness of the machine learning framework, but is in fact an advantage since twin nucleation and propagation are not sufficiently understood to create a model based on first principles. The idea of extracting structure/ phenomena relations, and ranking the importance of different structure attributes, without prior understanding of the physics, is perhaps the greatest benefit of using a machine learning approach. For example, if one considers a scenario where no previous work had been done to study twinning in Mg, then the results of machine learning could have led to focused studies on the highlighted individual attributes found in the decision trees to affect the phenomenon of twinning. Nonetheless, the findings from the machine learning framework have provided additional insights into deformation twinning in AZ31. These insights have the potential to guide future studies based upon first-principles.

In this case, some of the initial studies on particular attributes have already been performed, which helped to validate the results and framework presented in this paper. Because the attributes and values in the resulting decision trees are not limited to the physics encapsulated in previous models, they provide additional, unaccounted for insights into the attributes affecting twinning. One new attribute found using the machine learning framework is the deviation of the c-axis from the sample ND. An example of how this might lead to a more in-depth study would be to observe highly non-basal grains in a mostly basal textured material, and to come up with a measure of strain incompatibility that would then help explain twinning in these grains.

Another potential application of machine learning as it applies to the findings of this paper is the opportunity to refine and accelerate microscope-based studies on twinning in Mg, specifically AZ31. One implication of this is the ability to apply the results from this paper to quickly select areas of interest (e.g. regions where twin nucleation is probable) and capture high quality data in these areas. The model may then be applied to observe twin nucleation at its earliest detectable stages by looking for the highlighted attributes at lower levels of global strain. Such tools could serve to help future studies of any complex materials phenomenon observable by EBSD.

4.5 Summary

Twin nucleation within grains and twin propagation across grain boundaries in Mg alloy AZ31 was studied using a machine learning framework (J48 decision tree), by extracting attribute based rules. Models emerging from the study were 86.5% accurate for predicting twin nucleation in 104 grains and 96.1% accurate for predicting twin propagation across grain boundaries. The test cases were 75.1% for twin nucleation in 1239 grains and the twin propagation model was 75.5% accurate on 1127 GBs of interest. While the test cases were less accurate than the cases used to build the models the test sample had a significantly greater amount of events than the model sample. Additionally, the compression strain and grain size distribution was slightly different between the test and model samples which may have led to some of the discrepancies. Even with accuracies around 75% the test samples still demonstrated the ability of the decision tree model to apply to a dataset ten times bigger than the one used to create the model.

The important observations associated with twin nucleation on the grain level, in order of the greatest relevance, include:

- increased twin nucleation in larger grains (>7.2 μm)
- decreased twin nucleation in grains with maximum basal Schmid factors above ~0.29
- decreased twin nucleation in grains with average dislocation densities above 2.23e15 m^{-2}
- increased twin nucleation in grains where the angular deviation between c-axis and sample ND was less than 30.5°

The important observations associated with twin propagation across a grain boundary, in order of relevance include:

- increased twin propagation across GBs with a length greater than 10.05 μm
- increased twin propagation at GB misorientations <27.1°
- decreased twin propagation for GBs where the maximum basal Schmid factor exceeded 0.359
- increased twin propagation across GBs where the difference between the ratio of twin Schmid factor to <c+a> Schmid factor was less than or equal to 0.317
- decreased twin propagation across GBs with a difference between kernel average misorientation greater than 0.032°
- decreased twin propagation across GBs where the average angle between the GB and loading direction was less than or equal to 46.63°

The combination of these attributes provide the greatest overall information gain of the available features and help understand the physical processes associated with tensile twinning in AZ31. Furthermore, the findings of this paper stand in agreement with work done by previous researchers (grain size corresponds to Hall-Petch/Barnett models (Barnett, Keshavarz et al. 2004), boundary length corresponds to Beyerlein (Beyerlein, McCabe et al. 2011), and GB

misorientation corresponds to both Khosravani (Khosravani, Fullwood et al. Submitted 2012) and Beyerlein (Beyerlein, McCabe et al. 2011)) while building upon it with the addition of new insights just enumerated previously. These additional insights provide a deeper understanding of some of the previously hidden details affecting twinning in AZ31; leading to potentially more accurate models in future work.

In conjunction with the findings on twinning in this Mg alloy, a reliable machine learning framework for utilizing EBSD data has been developed to analyze a variety of complex phenomena. The benefits of this new technique are two-fold: i) underlying processes, or microstructural features, behind studied events can be quickly revealed, and ii) continued study of a phenomenon will be aided by the ability to quickly refine data collection to areas of interest in a microstructure. For example, in the case of twinning in AZ31, the decision tree presented in this paper could be applied to coarse EBSD scans in order to limit subsequent scans to only those grains predicted to nucleate twins. This offers the potential to capture twin nucleation at its earliest stages, allowing for direct observations, which are currently difficult to pin point. The framework presented in this study is not limited to twinning in Mg or its alloys but extends to the broader set of crystalline materials. Ultimately this technique has potential for analyzing complex phenomena rapidly, which will allow for acceleration in research progress on the links between microstructure and material properties.

5 CONCLUSIONS

In seeking to further characterize deformation twins in Mg, a method was developed to measure surface topography and detect cracks using EBSD images. The method of topography recovery resulted in reconstruction errors as low as 2%, or 1 μm, over a 500 μm length. The method was applied to a 3 μm x 3 μm area of twinned Tantalum which experienced topographic alterations. The reconstructed topography of Ta was in line with theoretical estimations of surface changes caused by twinning and also correlated with other local deformation mechanisms. Additionally, EBSD images were used to identify the presence of cracks in Nickel microstructures based on shadowing on the EBSD detector. Cracks of several sizes were measured on the Ni specimen which revealed that cracks as thin as 34 nm could potentially be measured. Not only was it shown that crack could be detected, but also the direction of the cracks was captured based on the position of the shadowing in the EBSD images. These new tools have the tremendous advantage of collecting crystal information simultaneously with surface structure.

The other EBSD based tool developed for this dissertation involved tracking individual Kikuchi band intensity to identify thin compression twins in Mg, which are often missed due to their size relative to the electron interaction volume. This tool took advantage of matching planes across a twin boundary to confidently identify undetected twins in Mg. Proof of concept was performed on known twins in Inconel 600, Tantalum, and Magnesium AZ31. The tool was then

used to check streaks of low image quality inside a grain that might be twins. By searching a few areas with potential twins Kikuchi band tracking was able to increase the amount of identified compression twins in the microstructure by 83%.

In order to take advantage of the new information available for twins in Mg, a new technique was used to uncover the driving forces behind deformation twinning. A machine learning framework was developed to leverage all of the data available from EBSD and use that to create a physics-based set of rules, or model, that describe twin nucleation and growth. The resultant models for nucleation and growth were found to be up to 86.5% and 96.1% accurate respectively while still maintaining accuracies of 75.1% and 75.5% for much larger datasets with slightly different strain and grain sizes. The first model pointed to grain size, basal Schmid factor, dislocation density, and deviation from the basal texture as the key attributes of twin nucleation while twin propagation across grain boundaries depended more on grain boundary length, orientation relative to loading, misorientation, along with the comparison of <c+a> Schmid factor to twinning Schmid factor, kernel average misorientation, and basal Schmid factor across the grain boundary. The new insights from these models provide guidance into where future research of deformation twinning in Mg should focus.

Not only do the tools developed by this research provide direction for future work on Mg alloys, but they can also be applied to any number of crystalline materials. Parameters of the background ellipse may be used to extract information about the density of a material or even magnetic domains. Additionally, tracking individual Kikuchi band intensity could be used to measure grain boundary inclination from two-dimensional data. In summary, this dissertation has expanded the abilities of EBSD to identify and characterize deformation twinning in Mg in

addition to building a framework where that data or other material data can be processed to gain further understanding of physical processes in materials.

REFERENCES

Adams, B.L., D.T. Fullwood, J. Basinger and T. Hardin. "Hr-Ebsd Microscopy for Localization Studies: Opportunities and Challenges." In *Int. Conf. Plasticity*. Puerto Vallarta, Mexico, 2011.

Adams, B.L., S.I. Wright and K. Kunze. "Orientation Imaging: The Emergence of a New Microscopy." *Metallurgical and Materials Transactions A* 24, no. 4 (1993): 819-831.

Adams, B.L., S.I. Wright and K. Kunze. "Orientation Imaging: The Emergence of a New Microscopy." *Metallurgical Transactions A (Physical Metallurgy and Materials Science)* 24A, no. 4 (1993): 819-31.

Adams, Brent L. "Orientation Imaging Microscopy: Emerging and Future Applications." *Ultramicroscopy* 67, no. 1-4 (1997): 11-17.

Agnew, S. R., M. H. Yoo and C. N. Tomé. "Application of Texture Simulation to Understanding Mechanical Behavior of Mg and Solid Solution Alloys Containing Li or Y." *Acta Materialia* 49, no. 20 (2001): 4277-4289.

Agnew, S.R., C.N. Tome, D.W. Brown, T.M. Holden and S.C. Vogel. "Study of Slip Mechanisms in a Magnesium Alloy by Neutron Diffraction and Modeling." *Scripta materialia* 48, (2003): 1003-1008.

Agnew, SR and O. Duygulu. "Plastic Anisotropy and the Role of Non-Basal Slip in Magnesium Alloy Az31b." *Int. J. Plasticity* 21, (2005): 1161-1193.

Al-Samman, T. and G. Gottstein. "Room Temperature Formability of a Magnesium Az31 Alloy: Examining the Role of Texture on the Deformation Mechanisms." *Materials Science and Engineering A* 488, (2008): 406-414.

Alam, M.N., M. Blackman and D.W. Pashley. "High-Angle Kikuchi Patterns." *Proc Roy Soc London* A221, (1954): 224-242.

Altinkok, N. and R. Koker. "Modelling of the Prediction of Tensile and Density Properties in Particle Reinforced Metal Matrix Composites by Using Neural Networks." *Materials & design* 27, no. 8 (2006): 625-631.

B.C. Wonziewiz , W.A. Backofen. "Plasticity of Magnesium Crystals." *Transactions of the Metallurgical Society of AIME* 239, (1967): 1422-1433.

Barnett, M.R. "Twinning and the Ductility of Magnesium Alloys: Part I: 'Tension' Twins." *Materials Science and Engineering: A* 464, no. 1-2 (2007): 1-7.

Barnett, M.R., Z. Keshavarz, A.G. Beer and D. Atwell. "Influence of Grain Size on the Compressive Deformation of Wrought Mg–3al–1zn " *Acta Materialia* 52, no. 17 (2004): 5093-5103.

Barnett, MR. "Twinning and the Ductility of Magnesium Alloys Part Ii Contraction Twins." *Mater. Sci. Eng A* 464, (2007): 8-16.

Basinger, Jay, David Fullwood and Brent Adams. "Ebsd Detail Extraction for Greater Spatial and Angular Resolution in Material Characterization." In *TMS*. San Diego, 2011.

Becker, R. "Effects of Strain Localization on Surface Roughening During Sheet Forming." *Acta Materialia* 46, no. 4 (1998): 1385-1401.

Behm, Th, C. Funke and H. J. Möller. "Surface Orientation Characterisation of Rough Mc-Silicon Surfaces by Confocal Microscopy and Ebsd." *Surface and Interface Analysis* 45, no. 4 (2013): 781-786.

Beyerlein, I J and C N Tome. "A Probabilistic Twin Nucleation Model for Hcp Polycrystalline Metals." *Proceedings of the Royal Society A* 466, (2010): 2517-2544.

Beyerlein, I.J., R.J.McCabe and C.N.Tome. "Effect of Microstructure on the Nucleation of Deformation Twins in Polycrystalline High-Purity Magnesium: A Multi-Scale Modeling Study." *Journal of the Mechanics and Physics of Solids* 59, (2011): 988-1003.

Beyerlein, IJ, RJ McCabe and CN Tomé. "Effect of Microstructure on the Nucleation of Deformation Twins in Polycrystalline High-Purity Magnesium: A Multi-Scale Modeling Study." *Journal of the Mechanics and Physics of Solids*, (2011).

Brewer, Luke N., David P. Field and Colin C. Merriman. "Mapping and Assessing Plastic Deformation Using Ebsd." In *Electron Backscatter Diffraction in Materials Science, 2nd Ed.*, edited by A.J. Schwartz, M. Kumar, B.L. Adams and D.P. Field, 251-262: Springer, 2009.

Britton, T.B. and A.J. Wilkinson. "Measurement of Residual Elastic Strain and Lattice Rotations with High Resolution Electron Backscatter Diffraction." *Ultramicroscopy* 111, no. 8 (2011): 1395-1404.

Capolungo, L., I. J. Beyerlein and C. N. Tome. "Slip-Assisted Twin Growth in Hexagonal Close-Packed Metals." *Scripta materialia* 60, no. 1 (2009): 32-35.

Chapuis, A. and J. H. Driver. "Temperature Dependency of Slip and Twinning in Plane Strain Compressed Magnesium Single Crystals." *Acta Materialia* 59, no. 5 (2011): 1986-1994.

Chapuis, Adrien and Julian Driver. "A Fundamental Study of the High Temperature Deformation Mechanisms of Magnesium." *Journal of Physics: Conference Series* 240, (2010): 012092.

Chen, D., G. T. Lui and J. C. Kuo. "Application of Edge Detection Method Based on Image Quality Gradient for Twin Detection." *J Microsc* 236, no. 1 (2009): 44-51.

Christian, John Wyrill and Subhash Mahajan. "Deformation Twinning." *Progress in materials science* 39, no. 1 (1995): 1-157.

Deal, A., X. Tao and A. Eades. "Ebsd Geometry in the Sem: Simulation and Representation." *Surface and interface analysis* 37, no. 11 (2005): 1017-1020.

Demir, E., D. Raabe, N. Zaafarani and S. Zaefferer. "Investigation of the Indentation Size Effect through the Measurement of the Geometrically Necessary Dislocations beneath Small Indents of Different Depths Using Ebsd Tomography." *Acta Materialia* 57, no. 2 (2009): 559-569.

Fallahi, A. and A. Ataee. "Effects of Crystal Orientation on Stress Distribution near the Triple Junction in a Tricrystal Γ-Tial." *Materials Science and Engineering: A* 527, no. 18-19 (2010): 4576-4581.

Field, D. "Recent Advances in the Application of Orientation Imaging." *Ultramicroscopy* 67, (1997): 1-9.

Fullwood, David, Brent Adams, Jay Basinger, Timothy Ruggles, Ali Khosravani, Caroline Sorensen and Joshua Kacher. "Microstructure Detail Extraction Via Ebsd: An Overview." In *Strains and Dislocation Gradients from Diffraction*, edited by R Barrabas and G. E. Ice: Imperial College Press, 2014.

Gardner, Calvin J., Brent L. Adams, John Basinger and David T. Fullwood. "Ebsd-Based Continuum Dislocation Microscopy." *International Journal of Plasticity* 26, (2010): 1234-1247.

Geiss, R.H., R.R. Keller and D.T. Read. "Transmission Electron Diffraction from Nanoparticles, Nanowires and Thin Films in an Sem with Conventional Ebsd Equipment." *Microsc Microanal* 16, no. Suppl 2 (2010): 1742-1743.

Goldstein, J., D. Newbury, D.C. Joy, C. Lyman, P. Echlin, C. Lifshin, L. Saywer and J. Michael. *Scanning Electron Microscopy and X-Ray Mircoanalysis*. New York: Kluwer Academic / Plenum Publishers, 2003.

Graff, Stephane, Wolfgang Brocks and Dirk Steglich. "Yielding of Magnesium: From Single Crystal to Polycrystalline Aggregates." *International Journal of Plasticity* 23, (2007): 2007.

Haj-Ali, R., H.K. Kim, S.W. Koh, A. Saxena and R. Tummala. "Nonlinear Constitutive Models from Nanoindentation Tests Using Artificial Neural Networks." *International Journal of Plasticity* 24, no. 3 (2008): 371-396.

Hjelen, J. and Nes, E. "Spatial Resolution Measurements of Electron Backscatter Diffraction Patterns (Ebsps) in the Scanning Electron Microscope." *Proceedings of the 12th International Congress on Electron Microscopy*, (1990): 404-405.

Humphreys, FJ. "Characterisation of Fine-Scale Microstructures by Electron Backscatter Diffraction (Ebsd)." *Scripta materialia* 51, no. 8 (2004): 771-776.

Oim 6.0. EDAX-TSL.

Izadbakhsh, A., K. Inal, R. K. Mishra and M. Niewczas. "New Crystal Plasticity Constitutive Model for Large Strain Deformation in Single Crystals of Magnesium." *Computational materials science* 50, no. 7 (2011): 2185-2202.

J Wang, I Beyerlein, C Tome. "An Atomisitic and Probabilistic Perspective on Twin Nucleation in Mg." *Scripta materialia* 63, (2010): 741-746.

J Wang, R Hoagland, J Hirth, L Capolungo, I Beyerlein, C Tome. "Nucleation of a (1012) Twin in Hexagonal Close-Packed Crystals." *Scripta materialia* 61, (2009): 903-906.

Jon Scott, Michael Miles, David Fullwood, Brent Adams, Ali Khosravani, Raja K. Mishra. "Room Temperature Shear Band Development in Highly Twinned Wtought Magnesium Az31b Sheet." *Metall. Mater. Trans.A* 44, no. A (2012): 512-516.

Jonas, John J., Sijia Mu, Talal Al-Samman, Günter Gottstein, Lan Jiang and Ėtienne Martin. "The Role of Strain Accommodation During the Variant Selection of Primary Twins in Magnesium." *Acta Materialia* 59, no. 5 (2011): 2046-2056.

Joy, D.C. *Monte Carlo Modeling for Electron Microscopy and Microanalysis*. Oxford: Oxford University Press, 1995.

Kacher, Josh, Brent Adams and David Fullwood. "Detection of Tetragonalities and Pseudo-Symmetries by High Resolution Ebsd Methods." In *MS&T*. Pittsburgh, 2009.

Kacher, Josh, Colin Landon, Brent L. Adams and David Fullwood. "Bragg's Law Diffraction Simulations for Electron Backscatter Diffraction Analysis." *Ultramicroscopy* 109, no. 9 (2009): 1148-1156.

Khosravani, A, J Scott, MP Miles, D Fullwood, BL Adams and RK Mishra. "Twinning in Magnesium Alloy Az31b under Different Strain Paths at Moderately Elevated Temperatures." *International Journal of Plasticity* 45, (2013): 160-173.

Khosravani, Ali, David Fullwood, John Scott, Michael Miles and Raj Mishra. "Nucleation and Propagation of 1012 Twins in Az31 Magnesium Alloy." *Acta Materialia*, (Submitted 2012).

Klimanek, P. and A. Pötzsch. "Microstructure Evolution under Compressive Plastic Deformation of Magnesium at Different Temperatures and Strain Rates." *Materials Science and Engineering: A* 324, no. 1–2 (2002): 145-150.

Koike, J, T Kobayashi, T Mukai, H Watanabe, M Suzuki, K Maruyama and K. Higashi. "The Activity of Non-Basal Slip Systems and Dynamic Recovery at Room Temperature in Fine-Grained Az31b Magnesium Alloys." *Acta Mater* 51, (2003): 2055-65.

Koike, J. "Enhanced Deformation Mechanisms by Anisotropic Plasticity in Polycrystalline Magnesium Alloys at Room Temperature." *Metall. Mater. Trans.A* 36, (2005): 1689-96.

Koker, R., N. Altinkok and A. Demir. "Neural Network Based Prediction of Mechanical Properties of Particulate Reinforced Metal Matrix Composites Using Various Training Algorithms." *Materials & design* 28, no. 2 (2007): 616-627.

Landon, C., B. Adams and J. Kacher. "High Resolution Methods for Characterizing Mesoscale Dislocation Structures." *Journal of Engineering Materials and Technology* 130, no. 2 (2008): 40-45.

Maitland, T.M. and A. Gholinia. "Advances in High-Speed Ebsd Orientation Mapping." *Microsc Microanal* 13, (2007): 924-925.

Marion Calcagnatto, Dirk Ponge, Eralp Demir, Dierk Raabe. "Orientation Gradients and Geometrically Necessary Dislocations in Ultrafine Grained Dual-Phase Steels Studied by 2d and 3d Ebsd." *Materials Science and Engineering: A* 527, (2010): 2738-2746.

Meng, L., P. Yang, Q. Xie and W. Mao. "Analyses on Compression Twins in Magnesium." *Materials Transactions* 49, no. 4 (2008): 710-714.

Meyers, M. A., O. Vohringer and V. A. Lubarda. "The Onset of Twinning in Metals: A Constitutive Description." *Acta Materialia* 49, no. 19 (2001): 4025-4039.

Mordike, B. L. and T. Ebert. "Magnesium: Properties — Applications — Potential." *Materials Science and Engineering: A* 302, no. 1 (2001): 37-45.

N Ono, R Nowak, S Miura. "Effect of Deformation Temperature on Hall-Petch Relationship Registered for Polycrystalline Magnesium." *Materials Letters* 58, (2003): 39-43.

Pérez-Benitez, JA and LR Padovese. "Feature Selection and Neural Network for Analysis of Microstructural Changes in Magnetic Materials." *Expert Systems with Applications*, (2011).

Piao, K., K. Chung, M. G. Lee and R. H. Wagoner. "Twinning-Slip Transitions in Mg Az31b." *Metallurgical and Materials Transactions a-Physical Metallurgy and Materials Science* 43A, no. 9 (2012): 3300-3313.

Prasad, Dilip K., Maylor K. H. Leung and Chai Quek. "Ellifit: An Unconstrained, Non-Iterative, Least Squares Based Geometric Ellipse Fitting Method." *Pattern Recognition* 46, no. 5 (2013): 1449-1465.

Quinlan, J.R. "C4.5: Programs for Machine Learning." *Morgan Kaufmann Publishers*, (1993).

Raabe, Dierk, Michael Sachtleber, Hasso Weiland, Georg Scheele and Zisu Zhao. "Grain-Scale Micromechanics of Polycrystal Surfaces During Plastic Straining." *Acta Materialia* 51, no. 6 (2003): 1539-1560.

Rasmussen, J. "Skill, Rule, and Knowledge: Signals, Signs, and Symbols and Other Distinctions in Human Performance Models." *IEEE Transaction on Systems, Man, and Cybernetics* 13, no. 3 (1983): 257-266.

Reddy, NS, AK Rao, M. Chakraborty and BS Murty. "Prediction of Grain Size of Al–7si Alloy by Neural Networks." *Materials Science and Engineering: A* 391, no. 1 (2005): 131-140.

Reich, Yoram and Nahum Travitzky. "Machine Learning of Material Behaviour Knowledge from Empirical Data." *Materials & design* 16, no. 5 (1995): 251-259.

Ren, S. X., E. A. Kenik, K. B. Alexander and A. Goyal. "Exploring Spatial Resolution in Electron Back-Scattered Diffraction Experiments Via Monte Carlo Simulation." *Microscopy and Microanalysis* 4, no. 01 (1998): 15-22.

Ren, S.X., E.A. Kenik, K.B. Alexander and A. Goyal. "Exploring Spatial Resolution in Electron Back-Scattered Diffraction Experiments Via Monte Carlo Simulation." *Microsc. Microanal.* 4, (1998): 15-22.

Schwartz, A.J., Kumar, M., Adams, B.L. Field, D. *Electron Backscatter Diffraction in Materials Science*: Springer US, 2009.

Scott, J., M.P. Miles, D. Fullwood, B.L. Adams, A. Khosravani and R.K. Mishra. "Development of Shear Bands with High Twin Density in Wrought Magnesium Az31b Sheet at Room Temperature." *Met Trans A* 44a, (2013): 512-516.

Sha, W. and KL Edwards. "The Use of Artificial Neural Networks in Materials Science Based Research." *Materials & design* 28, no. 6 (2007): 1747-1752.

Sitzman, S. D., G. Nolze and M. M. Nowell. "Ebsd Pattern Quality and Its Use in Evaluating Sample Surface Condition." *Microscopy and Microanalysis* 16, no. S2 (2010): 698-699.

Sorensen, Caroline, JohnA Basinger, MatthewM Nowell and DavidT Fullwood. "Five-Parameter Grain Boundary Inclination Recovery with Ebsd and Interaction Volume Models." *Metallurgical and Materials Transactions A* 45, no. 9 (2014): 4165-4172.

Steinmetz, D. and S. Zaefferer. "Towards Ultra-High Resolution Ebsd by Low Accelerating Voltage." *Materials Science and Technology* 26, (2010): 640-645.

Stoudt, Mark R. and Joseph B. Hubbard. "Fundamental Relationships between Deformation-Induced Surface Roughness, Critical Strain Localisation and Failure in Aa5754-O." *Philosophical Magazine* 89, no. 27 (2009): 2403-2425.

Tao, Xiaodong and Alwyn Eades. "Monte Carlo Simulation for Electron Backscattering Diffraction." *Microsc Microanal* 10, no. 2 (2004): 940-941.

Taylor, G. I. "Plastic Strain in Metals." *Journal of the Institute of Metals* 62, (1938): 307-324.

Matlab 7.3. The Mathworks, Inc.

Tompos, A., J.L. Margitfalvi, E. Tfirst and K. Héberger. "Predictive Performance of "Highly Complex" Artificial Neural Networks." *Applied Catalysis A: General* 324, (2007): 90-93.

Valkonen, A.E. "Plastic Deformation and Roughness of Free Metal Surfaces." The Ohio State University, 1987.

Vaudin, MD, G. Stan, YB Gerbig and RF Cook. "High Resolution Surface Morphology Measurements Using Ebsd Cross-Correlation Techniques and Afm." *Ultramicroscopy*, (2011).

Villert, S., C. Maurice, C. Wyon and R. Fortunier. "Accuracy Assessment of Elastic Strain Measurement by Ebsd." *Journal of Microscopy* 233, no. 2 (2009): 290–301.

Wagoner RH, Lou XY, Li M, Agnew SR;. "Forming Behavior of Magnesium Sheet." *J. Mater. Proc. Tech*, no. 177 (2006): 483-5.

Waikato Environment for Knowledge Analysis (Weka). The University of Waikato, Hamilton, New Zealand.

Wilkinson, A. J., G. Meaden and D. J. Dingley. "Elastic Strain Tensor Mapping - Extending the Limits of Ebsd Analysis." *Microsc Microanal* 11, (2005): 520-521.

Wilkinson, A. J., G. Meaden and D. J. Dingley. "High Resolution Mapping of Strains and Rotations Using Electron Back Scatter Diffraction." *Materials Science and Technology* 22, no. 11 (2006): 1-11.

Wilkinson, Angus J., Graham Meaden and David J. Dingley. "High-Resolution Elastic Strain Measurement from Electron Backscatter Diffraction Patterns: New Levels of Sensitivity." *Ultramicroscopy* 106, (2006): 307-313.

Wolfer, Marco, Jürgen Biener, Bassem S. El-dasher, Monika M. Biener, Alex V. Hamza, Armin Kriele and Christoph Wild. "Crystallographic Anisotropy of Growth and Etch Rates of Cvd Diamond." *Diamond and Related Materials* 18, no. 5–8 (2009): 713-717.

Wright, S. I. "A Parametric Study of Electron Backscatter Diffraction Based Grain Size Measurements." *Praktische Metallographie-Practical Metallography* 47, no. 1 (2010): 16-33.

Wright, S.I. and B.L. Adams. "Automated Lattice Orientation Determination from Electron Backscatter Kikuchi Diffraction Patterns." *Textures and Microstructures* 14-18, (1991): 273-278.

Wright, Stuart I. "Review of Automated Orientation Imaging Microscopy (Oim)." *Journal of Computer-Assisted Microscopy* 5, no. 3 (1993): 207-221.

Wright, Stuart I. and Matthew M. Nowell. "Ebsd Image Quality Mapping." *Microscopy and Microanalysis* 12, (2006): 72-84.

Wright, Stuart I., Matthew M. Nowell, René de Kloe, Patrick Camus and Travis Rampton. "Electron Imaging with an Ebsd Detector." *Ultramicroscopy* 148, no. 0 (2015): 132-145.

Xu, L., L. Wencong, J. Shengli, L. Yawei and C. Nianyi. "Support Vector Regression Applied to Materials Optimization of Sialon Ceramics." *Chemometrics and intelligent laboratory systems* 82, no. 1 (2006): 8-14.

Yang, Ping, Li Meng, Yisong Hu, Zude Zhao and Xueping Ren. "Analysis of Deformation Processes of Magnesium Alloy at Elevated Temperatures by Orientation Mapping." *Image Anal Stereol* 23, (2004): 53-61.

Yassar, R.S., O. AbuOmar, E. Hansen and M.F. Horstemeyer. "On Dislocation-Based Artificial Neural Network Modeling of Flow Stress." *Materials & design* 31, no. 8 (2010): 3683-3689.

Yoram, R. and T. Nahum. "Machine Learning of Material Behavior Knowledge from Empirical Data."

Yue, Jiansong Wan and Zhufeng. "Non-Uniform Stress Field and Stress Concentration Inducedby Grain Boundary and Triple Junction of Tricrystal." *J. Mater. Sci. Technol.* 18, no. 4 (2003): 303-308.

Zaefferer, S. "On the Formation Mechanisms, Spatial Resolution and Intensity of Backscatter Kikuchi Patterns." *Ultramicroscopy* 107, no. 2–3 (2007): 254-266.

Zaefferer, S. "Some Ideas on the Formation Mechanisms and Intensity Distribution of Backscatter Kikuchi Patterns." *Microc Microanal* 13, (2007): 928-929.

6 APPENDIX A

Electron reflection off of a tilted sample was used to check the effects of surface topography on the resultant background ellipse and its center positon on the EBSD detector. This appendix provides a visualization of the output of the simulation.

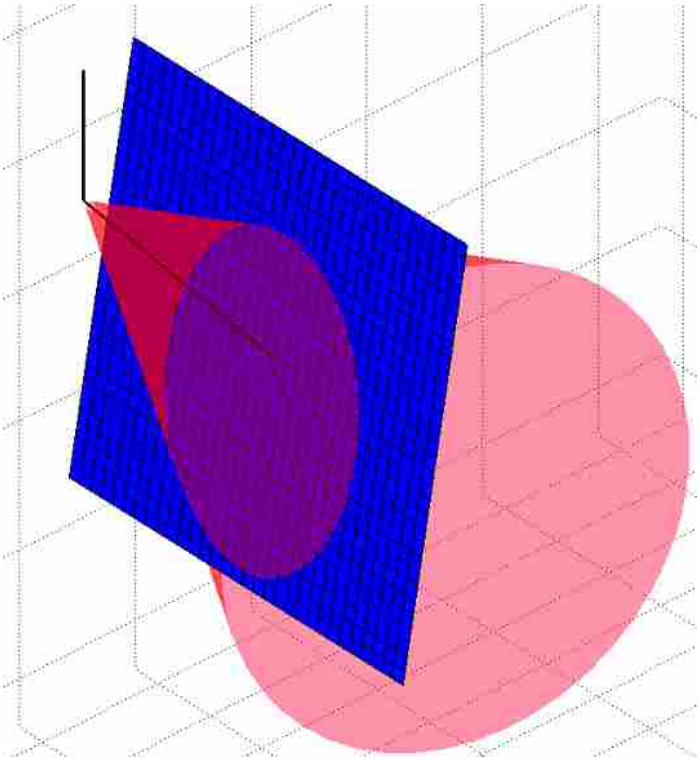


Figure 0-1: Graphical representation of Matlab simulations used to test effects of sample tilt on position of background ellipse center.

7 APPENDIX B

This section contains Matlab code developed to process EBSD information for input into creating a machine learning model for twin nucleation:

```
% Program to create Input for Machine Learning on Twins

% Looks at grains
% Travis Rampton
% 6/22/2011

more=1;

while more==1
% get input files
qstring = 'Use new data?';
choice = questdlg(qstring,'Another file or not','New data','Old
data','Old data');
switch choice
   case 'New data'
      clc
      clear all
      close all
      new=1;
      [name path]=uigetfile('*.mat','Select .mat file of data files'); %
Create with MLInputFiles.m
      load([path name])
      if strcmp(lattice,'hexagonal')
         lattice=2;
      else
         lattice=1;
      end
   case 'Old data'

      new=0;
end
format compact
%% Use .txt files

% Select custom grain file (.txt) to use
% fid=fopen([testpath testname]);
gvals=ReadCustGrainFile(testpath, testname);
data.ID=gvals{1};
data.phis(:,1)=gvals{2};%radians
data.phis(:,2)=gvals{3};%radians
data.phis(:,3)=gvals{4};%radians
data.xy(:,1)=gvals{5};
data.xy(:,2)=gvals{6};
data.IQ=gvals{7};
data.CI=gvals{8};
data.Fit=gvals{9};
data.edge=gvals{10};
data.numpts=gvals{11};
data.size=gvals{12};
data.numnbrs=gvals{13};
for i=1:max(data.numnbrs)
   data.nbrs(:,i)=gvals{13+i};
end
data.nbrs(isnan(data.nbrs))=0;
clear gvals

%% Scan standard file for data

% Select standard grain file (.txt) for use
ginputd=cell2mat(ReadStdGrainFile(testpath2,testname2));

%% Continue processing data
% Still needed: Avg nbr size / Ave nbr miso / # CSL boundaries /
% Miso from ND / Schmid / Taylor / Twin present

% Misorientation of c-axis to sample axes

rdMiso=imread([rdpath rdname]);
rdMisodata=rgb2gray(rdMiso);
rdMisodata=double(rdMisodata);

tdMiso=imread([tdpath tdname]);
tdMisodata=rgb2gray(tdMiso);
tdMisodata=double(tdMisodata);

ndMiso=imread([ndpath ndname]);
ndMisodata=rgb2gray(ndMiso);
ndMisodata=double(ndMisodata);

% Create grain ID point by point map
step=abs(abs(ginputd(2,4))-abs(ginputd(1,4))); % Careful with this
step, needs to be more robust
[numy numx]=size(rdMisodata);
grainID=zeros(size(rdMisodata));
for i=1:numy*numx

grainID(round(ginputd(i,5)/step)+1,round(ginputd(i,4)/step)+1)=gi
nputd(i,9);
end

% mmaxtemp=[myname(12:13),'.',myname(15:18)];
mimax=90;
mirange=mimax;

rdMisodata=(rdMisodata*mirange/255);
tdMisodata=(tdMisodata*mirange/255);
```

```
ndMisodata=(ndMisodata*mirange/255);
% figure;pcolor(flipud(Misodata)); shading flat; colorbar

% Read in twin data
% uigetfile({'*.jpg;*.tif;*.png;*.bmp','All Image Files';'*.*','All
Files' },'Select twin parent image file',testpath);
twinparent=imread([tpath tname]);
tpinfo=rgb2gray(twinparent);
% untwinned = 255
% parent = 29
% daughter = 76
% Find area that is not twins
nottwins=ones(size(rdMisodata));
nottwins(tpinfo==76)=0;% Initial twin assignment

% figure;pcolor(flipud(grainID));shading flat
% Last row of data has been removed from image
% Add grain boundary points
gbpts=zeros(size(grainID));
for n=2:length(grainID(:,1))-1
   for m=2:length(grainID(1,:))-1
      numtpts=0;
      for a=n-1:n+1
         for b=m-1:m+1
            if grainID(n,m)~=grainID(a,b)
               numtpts=numtpts+1;
            end
         end
      end
      if numtpts>2
         gbpts(n,m)=1;
      end
   end
end
% gbpts=bwmorph(gbpts,'skel',Inf);
[ygbs xgbs]=find(gbpts==1);
r=a;
c=b;

phi1=reshape(ginputd(:,1),[c r])';
PHI=reshape(ginputd(:,2),[c r])';
phi2=reshape(ginputd(:,3),[c r])';

figure('Name','Original vs Changed twin map','NumberTitle','off');
subplot(1,2,1);imagesc(nottwins);
hold on
plot(xgbs,ygbs,'k*','MarkerSize',2)
axis equal tight
% adjustment for twin dominated grain
% check misorientation from 0001
changednt=zeros(size(data.ID));

if btexture(1)=='Y'

   lnot=zeros(size(data.ID));
   ltwin=zeros(size(data.ID));
   avenot=zeros(size(data.ID));
   avetwin=zeros(size(data.ID));
   mtwin=zeros(size(data.ID));
   ndtwin=ndMisodata.*(1-nottwins);
   ndnot=ndMisodata.*nottwins;
   for k=1:length(data.ID)
      lnot(k)=sum(ndnot(grainID==k)>0);
      ltwin(k)=sum(ndtwin(grainID==k)>0);
      avenot(k)=sum(ndnot(grainID==k))/lnot(k);
      avetwin(k)=sum(ndtwin(grainID==k))/ltwin(k);
      mtwin(k)=mean(ndtwin(grainID==k));
      if avenot(k)>2*avetwin(k) && mtwin(k)~=0
         disp(['grain ',num2str(k)])
         ndtemp=ndMisodata;
         ndtemp(grainID~=data.ID(k))=0;
         figure(2);subplot(1,2,1);imagesc(ndtemp);title('Pick point
for average Euler angles (Click left of map to not change)');
         caxis([0 90]);axis equal tight
         nottwinsg=nottwins;
         nottwinsg(ndtemp==0)=0;
         subplot(1,2,2);imagesc(nottwinsg);title('red is untwinned
area');axis equal tight
         set(gcf, 'Position', get(0,'Screensize')); % Maximize figure.
         [cx ry]=ginput(1);
         close 2
         cx=round(cx);ry=round(ry);
         if cx>0
            data.phis(i,:)=[phi1(ry,cx),PHI(ry,cx),phi2(ry,cx)];
            nottwins(grainID==k)=1-nottwins(grainID==k);%
Switch parent and daughter
            changednt(k)=1;
         end
      end

   end
end
% figure;
subplot(1,2,2);imagesc(nottwins);
hold on
plot(xgbs,ygbs,'k*','MarkerSize',2)
axis equal tight

% subplot(1,2,2);imagesc(grainID);axis equal tight

for i=1:length(data.ID)
   if changednt(i)==1


   end
end


data.rdmiso=zeros(size(data.ID));
data.tdmiso=zeros(size(data.ID));
data.ndmiso=zeros(size(data.ID));
for k=1:length(data.ID)
   rdtemp=rdMisodata;
   rdtemp(grainID~=data.ID(k))=0;
   tdtemp=tdMisodata;
   tdtemp(grainID~=data.ID(k))=0;
   ndtemp=ndMisodata;
   ndtemp(grainID~=data.ID(k))=0;

data.rdmiso(k,1)=sum(sum(rdtemp.*nottwins))/sum(sum((rdtemp.
*nottwins)>0));

data.tdmiso(k,1)=sum(sum(tdtemp.*nottwins))/sum(sum((tdtemp.*
nottwins)>0));

data.ndmiso(k,1)=sum(sum(ndtemp.*nottwins))/sum(sum((ndtemp
.*nottwins)>0));
end


for i=1:length(data.ID)
   A=euler2gmat(data.phis(i,1),data.phis(i,2),data.phis(i,3));
   nbrsizes=0;
   nbrmiso=0;
   nbrmisolist=zeros(1,length(data.nbrs(1,:)));
   for j=1:length(data.nbrs(1,:))
      if data.nbrs(i,j)~=0 && ~isnan(data.nbrs(i,j))
         % Average neighbor sizes
         nind=find(data.ID==data.nbrs(i,j));
```

```
        nbrsizes=nbrsizes+data.size(nind);
        % Neighbor misorientations

B=euler2gmat(data.phis(nind,1),data.phis(nind,2),data.phis(nind,3)
);
        curmiso=GeneralMisoCalc(A,B,lattice);
        nbrmiso=nbrmiso+curmiso;
        nbrmisolist(j)=curmiso;
      end
   end
   nbrmisolist(nbrmisolist==0)=[];
   data.nbrmisomin(i,1)=min(nbrmisolist);
   data.nbrmisomax(i,1)=max(nbrmisolist);
   data.nbrsizes(i,1)=nbrsizes/data.numnbrs(i);
   data.nbrmiso(i,1)=nbrmiso/data.numnbrs(i);

end

% Find grains that contain twins
data.twins=zeros(size(data.ID));
for i=1:length(data.ID)
   if length(unique(tpinfo(grainID==data.ID(i))))>1
      data.twins(i)=1;
   end
end
% Check to see if correct
twincheck1=zeros(size(tpinfo));
twincheck1(tpinfo==29)=1;
twincheck1(tpinfo==76)=1;
figure('Name','Twinned Grains
(red)','NumberTitle','off');imagesc((twincheck1)); shading flat
hold on
plot(xgbs,ygbs,'*k','MarkerSize',2)

twincheck2=zeros(size(tpinfo));
for i=1:numy
   for j=1:numx
      if data.twins(find(grainID(i,j)))==1
         twincheck2(i,j)=1;
      end
   end
end
% figure;pcolor(flipud(twincheck2)); shading flat

tc=twincheck2-twincheck1;
% figure; pcolor(flipud(tc)); shading flat
% Appears to be small discrepancy between 2 twin methods

% Read in Schmid and Taylor factor images for conversion
% Requires reading in parent data to exclude info from twin
% % Schmid

load([spath2 sname2]);% load Schmid Factor data from code

schmid=imread([spath sname]);
schmiddata=rgb2gray(schmid);
imschmiddata=double(schmiddata);


smin=0;
smax=0.5;
srange=smax-smin;
imschmiddata=imschmiddata*smax/255;
schmiddata=sfbmap;
if ~min(size(nottwins)==size(schmiddata))
   disp('Schmid Factor file different size than other files')
   schmiddata(end,:)=[];%take out last row since this is the one
missing from other data
end
data.schmidb=zeros(size(data.ID));
for k=1:length(data.ID)
   stemp=schmiddata;
   stemp(grainID~=k)=0;

data.schmidb(k,1)=sum(sum(stemp.*nottwins))/sum(sum((stemp.*
nottwins)>0));
end

schmid1=imread([spath1 sname1]);
schmiddata1=rgb2gray(schmid1);
imschmiddata1=double(schmiddata1)*smax/255;
schmiddata1=sfcamap;
if ~min(size(nottwins)==size(schmiddata1))
   disp('Schmid Factor file different size than other files')
   schmiddata1(end,:)=[];%take out last row since this is the one
missing from other data
end
data.schmidca=zeros(size(data.ID));
for k=1:length(data.ID)
   stemp=schmiddata1;
   stemp(grainID~=k)=0;

data.schmidca(k,1)=sum(sum(stemp.*nottwins))/sum(sum((stemp.
*nottwins)>0));
end
figure('Name','Twin Schmid Factor','NumberTitle','off');
subplot(1,2,1);imagesc(imschmiddata1);axis equal tight
title('image file');
caxis([smin smax])

schmiddata2=sftwinmap;
if ~min(size(nottwins)==size(schmiddata2))
   disp('Schmid Factor file different size than other files')
   schmiddata2(end,:)=[];%take out last row since this is the one
missing from other data
end
% schmiddata2=rgb2gray(schmid2);
% schmiddata2=double(schmiddata2)*smax/255;
data.schmidt=zeros(size(data.ID));
for k=1:length(data.ID)
   stemp=schmiddata2;
   stemp(grainID~=k)=0;

data.schmidt(k,1)=sum(sum(stemp.*nottwins))/sum(sum((stemp.*
nottwins)~=0));
end
subplot(1,2,2);imagesc(schmiddata2);axis equal tight
title('mat file')
caxis([smin smax])

figure('Name','Basal Schmid Factor','NumberTitle','off');
subplot(1,2,1);imagesc(imschmiddata);axis equal tight
title('image file')
caxis([smin smax])
subplot(1,2,2);imagesc(sfbmap);axis equal tight
title('mat file')
caxis([smin smax])

figure('Name','c+a Schmid Factor','NumberTitle','off');
subplot(1,2,1);imagesc(imschmiddata1);axis equal tight
title('image file')
caxis([smin smax])
subplot(1,2,2);imagesc(sfcamap);axis equal tight
title('mat file')
caxis([smin smax])


% Calculate self misorientation

Miso=imread([mypath myname]);
```

```
Misodata=rgb2gray(Miso);
Misodata=double(Misodata);

% mmaxtemp=[myname(12:13),'.',myname(15:18)];
mmax=5;
mrange=mmax;

Misodata=(Misodata*mrange/255);
% figure;pcolor(flipud(Misodata)); shading flat; colorbar

data.Misograin=zeros(size(data.ID));
for k=1:length(data.ID)
   mtemp=Misodata;
   mtemp(grainID~=k)=0;

data.Misograin(k,1)=sum(sum(mtemp.*nottwins))/sum(sum((mte
mp.*nottwins)>0));

   data.twinfraction(k,1)=1-
(sum(sum(nottwins(grainID==k)))/sum(sum(grainID==k)));
end

if DD
   load([ddpath ddname]);
   dd=reshape(alpha_data.alpha_total3,[Settings.data.cols
Settings.data.rows])';
   if ~min(size(nottwins)==size(dd))
      disp('Schmid Factor file different size than other files')
      dd(end,:)=[];%take out last row since this is the one missing
from other data
   end
   for k=1:length(data.ID)
      dtemp=dd;
      dtemp(grainID~=k)=0;

data.Disloc(k,1)=sum(sum(dtemp.*nottwins))/sum(sum((dtemp.*n
ottwins)>0));
      if isnan(data.Disloc(k,1))
         data.Disloc(k,1)=1;% equals 0 after log10
      end
   end

else
   dd=ones(size(grainID)); % default for no data
   for k=1:length(data.ID)
      dtemp=dd;
      dtemp(grainID~=k)=0;

data.Disloc(k,1)=sum(sum(dtemp.*nottwins))/sum(sum((dtemp.*n
ottwins)>0));

   end
end
data.Disloc=log10(data.Disloc);

%% Save final file for input to machine learning

mlinput=zeros(length(data.ID),9+1+1); % 9 attributes and 1 twin
column and 1 column for edge grains

% Input Order
% 1 Size, 2 # of neighbors, 3 avg nbr size, 4 avg nbr miso, 5 avg
self miso, 6 miso
% from ND, 7 Taylor factor, 8 IQ, 9 CI, 10 TWIN present

mlinput(:,1)=data.size; % Reliable from custom grain file
mlinput(:,2)=data.numnbrs; % Reliable from custom grain file
mlinput(:,3)=data.nbrsizes; % Reliable from custom grain file
mlinput(:,4)=data.size./(data.nbrsizes./data.numnbrs); % relative
size of grain compared to neighbors * think about adding max and
min
mlinput(:,5)=data.nbrmiso;
mlinput(:,6)=data.nbrmisomin;
mlinput(:,7)=data.nbrmisomax;
mlinput(:,8)=data.Misograin; % Reliable from image analysis
mlinput(:,9)=data.ndmiso;
% mlinput(:,7)=data.taylorgrain; % Reliable from image analysis
mlinput(:,10)=data.schmidb; % Reliable from image analysis
mlinput(:,11)=data.schmidca; % Reliable from image analysis
mlinput(:,12)=data.schmidt; % Reliable from image analysis
mlinput(:,13)=data.schmidt./data.schmidca;
mlinput(:,14)=data.Disloc;
mlinput(:,15)=data.twins; % Reliable from image analysis
mlinput(:,16)=data.twinfraction;
mlinput(:,17)=-1*data.edge; % Reliable from custom grain file
mlinput(:,18)=data.ID; % Reliable from custom grain file

labels={'SIZE ' 'NUMNBRS ' 'NBRSIZE ' 'RELSIZE' 'NBRMISO '
'NBRMISOMIN ' 'NBRMISOMAX' 'MISOGRAIN ' 'NDMISO '
'BASALSF ' 'CASF ' 'TWINSF ' 'TWIN2CA' 'LOGDD ' 'TWIN01 '
'FRACTION ' 'EDGE ' 'ID '};

% Add in percent twinned

%Only use grains in the neighborhood of twins
%twinned grains
% twinnedgs=data.ID(data.twins==1);
% %grains touching grains with twins
% twinnbrs=data.nbrs(data.twins==1,:);
% twinnbrs=unique(twinnbrs);
% twinnbrs(1)=[];
% %grains touching grains that touch twinned grains
% for nbrhood=1:length(twinnbrs)
%    othernbrs=data.nbrs(data.ID==twinnbrs(nbrhood),:);
%    othernbrs=unique(othernbrs);
%    othernbrs(1)=[];
%    if nbrhood==1
%       newnbrs=othernbrs;
%    else
%       newnbrs=[newnbrs,othernbrs];
%       newnbrs=unique(newnbrs);
%    end
% end
% newnbrs=newnbrs';
UseIDs=unique(grainID);
% UseIDs=[twinnedgs;twinnbrs];
% UseIDs=[UseIDs;newnbrs];%Comment out if only direct
neighbors wanted
UseIDs=unique(UseIDs);
TrainIDs=zeros(size(UseIDs));
for use=1:length(UseIDs)
   TrainIDs(use)=find(data.ID==UseIDs(use));
end
mlinput2=mlinput(TrainIDs,:);
%Remove edge grains
% mlinput2=sortrows(mlinput2,size(mlinput,2)-1);
% mlinput2(1:abs(sum(mlinput2(:,size(mlinput,2)-1))),:)=[];
%Remove grain IDs
% mlinput2(:,size(mlinput2,2))=[];
rowcell=ones(length(mlinput2(:,1)+1),1);
colcell=ones(length(mlinput2(1,:)),1);

gbcell=mat2cell(mlinput2,rowcell,colcell);
% Save new file or append to already existent file
treetest=0;
qstring = 'Create new file OR append to existing file?';
choice = questdlg(qstring,'Next Action','New File','Append to
File','Test on Tree','New File');
```

```
switch choice
   case 'New File'
      % Save new file
      [filename, pathname] = uiputfile('*.xls','Save data as a csv
file',['C:\Users\Travis\Desktop\MLtwinsNickel\' testname(1:4)]);
      xlswrite([pathname filename],[labels;gbcell]);
   case 'Append to File'
      % Append to old file
      [filename, pathname] = uigetfile('*.xls','Append data to a
previous csv file','F:\RolledNickel\');
      xlssize=size(xlsread([pathname filename],1));
      endcol=num2letter(xlssize(2));
      xlswrite([pathname
filename],gbcell,['A',num2str(xlssize(1)+1),':',endcol,num2str(xlssi
ze(1)+size(gbcell,1))]);
   case 'Test on Tree'
      treetest=1;
end

%% Machine Learning Model for Twin Prediction taken from
WEKA
if treetest==1

% labels=['SIZE' 'NUMNBRS' 'NBRSIZE' 'NBRMISO'
'MISOGRAIN' 'NDMISO'...
% 'BASALSF' 'CASF' 'TWINSF' 'IQ' 'CI' 'RDMISO' 'TDMISO'
'LOGDD' 'TWIN01'...
% 'FRACTION' 'EDGE' 'ID'];

SIZE=mlinput2(:,1);
NumNbrs=mlinput2(:,2);
NbrSize=mlinput2(:,3);
NBRMISO=mlinput2(:,4);
MISOGRAIN=mlinput2(:,5);
NDMISO=mlinput2(:,6);
BASALSF=mlinput2(:,7);
CASF=mlinput2(:,8);
TWINSF=mlinput2(:,9);
IQ=mlinput2(:,10);
CI=mlinput2(:,11);
RDMiso=mlinput2(:,12);
TDMiso=mlinput2(:,13);
LOGDD=mlinput2(:,14);
Twin01=mlinput2(:,15);
Fraction=mlinput2(:,16);
Edge=mlinput2(:,17);
ID=mlinput2(:,18);

ptwin=zeros(size(mlinput2(:,1)));

for i=1:length(mlinput2(:,1))
   % change below text by running DTreader and pasting resulting
file
% 1 for twin and 0 for no twin
   if SIZE(i) <= 7.2
      ptwin(i)=0;
   else
      if BASALSF(i) <= 0.292391
         if LOGDD(i) <= 15.347673
            ptwin(i)=1;
         else
            ptwin(i)=0;
         end
      else
         if SIZE(i) <= 9.66
            ptwin(i)=0;
         else
            if NDMISO(i) <= 30.455882
               ptwin(i)=1;
            else
               ptwin(i)=0;
            end
         end
      end
   end


end

ptwins=ptwin;
% ptwins(isnan(ptwins))=0;
badpts=ptwins-mlinput2(:,15);
numbadpts=sum(abs(badpts));

missed=mlinput2(abs(badpts)==1,end);
falsetwin=mlinput2(badpts==1,end);
missedtwin=mlinput2(badpts==-1,end);

badgrains=zeros(size(tpinfo));
for i=1:numbadpts
   badgrains(grainID==missed(i))=1;
end
figure;imshow(badgrains);
hold on
plot(xgbs,ygbs,'*r','MarkerSize',2)
title('False Overall Predictions');

badtwins=ones(size(tpinfo))*255;
for i=1:length(falsetwin)
   badtwins(grainID==falsetwin(i))=rand()*100;
end
figure;imshow(uint8(badtwins));
hold on
plot(xgbs,ygbs,'*r','MarkerSize',1)
title('False Twin Prediction');

badnontwins=ones(size(tpinfo))*255;
for i=1:length(missedtwin)
   badnontwins(grainID==missedtwin(i))=rand()*100;
end
figure;imagesc(badnontwins);
hold on
plot(xgbs,ygbs,'*r','MarkerSize',1)
title('Missed Twin Prediction');

predtwins=zeros(size(tpinfo));
for i=1:length(ptwin)
   predtwins(grainID==mlinput2(i,end))=ptwin(i);
end
predtwins(predtwins==0)=0;
figure;imagesc(predtwins);
hold on
title('Twin Predictions by Leaf');
plot(xgbs,ygbs,'*r','MarkerSize',2)

% Plot correct and incorrect grains by color
Map=zeros(numy,numx,3);
Map(:,:,3)=255;
temp=zeros(numy,numx);
temp(badtwins<255)=255;
Map(:,:,1)=temp;
temp=temp*0;
temp(badnontwins<255)=255;
Map(:,:,2)=temp;
temp=temp*0;
temp(badgrains==0)=255;
Map(:,:,3)=temp;
temp=ones(numy,numx);
for i=1:length(data.edge)
   if data.edge(i)==1
```

```
        temp(grainID==mlinput(i,end))=.5;
    end
end
m1=Map(:,:,1);
m1(temp==.5)=125;
m2=Map(:,:,2);
m2(temp==.5)=125;
m3=Map(:,:,3);
m3(temp==.5)=125;
Map(:,:,1)=m1;
Map(:,:,2)=m2;
Map(:,:,3)=m3;

figure;imshow(uint8(Map));
shading flat; %colormapeditor;
hold on
title('Predictions');
plot(xgbs,ygbs,'*k','MarkerSize',2)

en
```

```
end
%% Decide to analyze another file or not

qstring = 'Analyze another file OR finish?';
choice = questdlg(qstring,'Another another file or not','Analyze a
New File','Finish','Finish');
switch choice
    case 'Analyze a New File'
        % Analyze another file
        close all
        clear all
        clc
        more=1;
    case 'Finish'
        % End file analysis
        more=0;
end
```

8 APPENDIX C

This section contains Matlab code developed to process EBSD information for input into creating a machine learning model for twin propagation. Due to length the code has been divided into two columns:

```
% Program to create Input for Machine Learning on Twins
% Looks at grain boundaries
% Travis Rampton
% 6/22/2011

more=1;

while more==1
%% Use .txt files
qstring = 'Use new data?';
choice = questdlg(qstring,'Another file or not','New data','Old
data','Old data');
switch choice
   case 'New data'
      clc
      clear all
      close all
      new=1;
      [name path]=uigetfile('*.mat','Select .mat file of data files'); %
Create with MLInputFiles.m
      load([path name])
      if strcmp(lattice,'hexagonal')
         lattice=2;
      else
         lattice=1;
      end
   case 'Old data'

      new=0;
end


if new==1
% Select custom grain file (.txt) to use
% [testname testpath]=uigetfile('*.txt','Custom Grain
File','F:\RolledNickel\Transfers\');
fid=fopen([testpath testname]);

% Read in and skip header lines
curline=fgetl(fid);
while numel(curline)==0
   curline=fgetl(fid);
end

while curline(1)=='#'
   curline=fgetl(fid);
   while numel(curline)==0
      curline=fgetl(fid);
   end
end

% Record first row of data
count=1;

format='%f %f %f %f %f %f %f %f %f %f %f %f %f %f';
curline2=textscan(curline,format,1);
for i=1:cell2mat(curline2(13))
   format=[format ' %f'];
end
curline2=textscan(curline,format,1);

data.ID(count,1)=cell2mat(curline2(1));
data.phis(count,:)=cell2mat(curline2(2:4));%be careful of this if
the input is radians or degrees
data.xy(count,:)=cell2mat(curline2(:,5:6));
data.IQ(count,1)=cell2mat(curline2(:,7));
data.CI(count,1)=cell2mat(curline2(:,8));
data.Fit(count,1)=cell2mat(curline2(:,9));
data.edge(count,1)=cell2mat(curline2(:,10));
data.numpts(count,1)=cell2mat(curline2(:,11));
data.size(count,1)=cell2mat(curline2(:,12));
data.numnbrs(count,1)=cell2mat(curline2(:,13));
if data.numnbrs(count,1)==0
   data.nbrs(count,1)=0;
else
   data.nbrs(count,1:i)=cell2mat(curline2(:,14:end));
end

% Record remaining rows of data

while ~feof(fid)

   curline=fgetl(fid);
   count=count+1;
   format='%f %f %f %f %f %f %f %f %f %f %f %f %f %f';
   curline2=textscan(curline,format,1);
   for i=1:cell2mat(curline2(13))
      format=[format ' %f'];
   end
   curline2=textscan(curline,format,1);
```

```
    data.ID(count,1)=cell2mat(curline2(1));
    data.phis(count,:)=cell2mat(curline2(2:4)); %be careful of this if
the input is radians or degrees
    data.xy(count,:)=cell2mat(curline2(:,5:6));
    data.IQ(count,1)=cell2mat(curline2(:,7));
    data.CI(count,1)=cell2mat(curline2(:,8));
    data.Fit(count,1)=cell2mat(curline2(:,9));
    data.edge(count,1)=cell2mat(curline2(:,10));
    data.numpts(count,1)=cell2mat(curline2(:,11));
    data.size(count,1)=cell2mat(curline2(:,12));
    data.numnbrs(count,1)=cell2mat(curline2(:,13));
    if data.numnbrs(count,1)==0
        data.nbrs(count,1)=0;
    else
        data.nbrs(count,1:i)=cell2mat(curline2(:,14:end));
    end

end
fclose(fid);

%% Scan standard file for data

% Select standard grain file (.txt) for use
% [testname2 testpath2]=uigetfile('*.txt','Standard Grain
File',testpath);
fid2=fopen([testpath2 testname2]);

% Read and skip header data
curline=fgetl(fid2);
while numel(curline)==0
    curline=fgetl(fid2);
end

while curline(1)=='#'
    curline=fgetl(fid2);
    while numel(curline)==0
        curline=fgetl(fid2);
    end
end

% Record first row of data
ginputd=zeros(sum(data.numpts),10);
count=1;

format='%f %f %f %f %f %f %f %f %f %f';
curline2=textscan(curline,format,1);

ginputd(count,:)=cell2mat(curline2);

% Record remaining rows of data
while ~feof(fid2)
    curline=fgetl(fid2);
    count=count+1;
    curline2=textscan(curline,format,1);
    ginputd(count,:)=cell2mat(curline2);
end

fclose(fid2);

%% Continue processing data
% Still needed: Avg nbr size / Ave nbr miso / # CSL boundaries /
% Miso from ND / Schmid / Taylor / Twin present

% Misorientation of c-axis to sample axes

% [rdname rdpath]=uigetfile({'*.jpg;*.tif;*.png;*.bmp','All Image
Files';...
%          '*.*','All Files' },'Select RD Miso image file',...
%           testpath);

rdMiso=imread([rdpath rdname]);
rdMisodata=rgb2gray(rdMiso);
rdMisodata=double(rdMisodata);

tdMiso=imread([tdpath tdname]);
tdMisodata=rgb2gray(tdMiso);
tdMisodata=double(tdMisodata);

ndMiso=imread([ndpath ndname]);
ndMisodata=rgb2gray(ndMiso);
ndMisodata=double(ndMisodata);

% Create grain ID point by point map
step=abs(abs(ginputd(2,4))-abs(ginputd(1,4))); % Careful with this
step, needs to be more robust
[numy numx]=size(rdMisodata);
grainID=zeros(size(rdMisodata));
for i=1:numy*numx

grainID(round(ginputd(i,5)/step)+1,round(ginputd(i,4)/step)+1)=gi
nputd(i,9);
end

% mmaxtemp=[myname(12:13),'.',myname(15:18)];
mimax=90;
mirange=mimax;

rdMisodata=(rdMisodata*mirange/255);
tdMisodata=(tdMisodata*mirange/255);
ndMisodata=(ndMisodata*mirange/255);
% figure;pcolor(flipud(Misodata)); shading flat; colorbar

% Read in twin data
% uigetfile({'*.jpg;*.tif;*.png;*.bmp','All Image Files';'*.*','All
Files' },'Select twin parent image file',testpath);
twinparent=imread([tpath tname]);
tpinfo=rgb2gray(twinparent);
% untwinned = 255
% parent = 29
% daughter = 76
% Find area that is not twins
nottwins=ones(size(rdMisodata));
nottwins(tpinfo==76)=0;% Initial twin assignment

% figure;pcolor(flipud(grainID));shading flat
% Last row of data has been removed from image
% Add grain boundary points
gbpts=zeros(size(grainID));
for n=2:length(grainID(:,1))-1
    for m=2:length(grainID(1,:))-1
        numtpts=0;
        for a=n-1:n+1
            for b=m-1:m+1
                if grainID(n,m)~=grainID(a,b)
                    numtpts=numtpts+1;
                end
            end
        end
        if numtpts>2
            gbpts(n,m)=1;
        end
    end
end
% gbpts=bwmorph(gbpts,'skel',Inf);
[ygbs xgbs]=find(gbpts==1);
r=a;
c=b;

phi1=reshape(ginputd(:,1),[c r])';
PHI=reshape(ginputd(:,2),[c r])';
```

```
phi2=reshape(ginputd(:,3),[c r])';

figure;subplot(1,2,1);imagesc(nottwins);
hold on
plot(xgbs,ygbs,'k*','MarkerSize',2)
axis equal tight
% adjustment for twin dominated grain
% check misorientation from 0001
changednt=zeros(size(data.ID));
temptwin=tpinfo;
if btexture(1)=='Y'

   lnot=zeros(size(data.ID));
   ltwin=zeros(size(data.ID));
   avenot=zeros(size(data.ID));
   avetwin=zeros(size(data.ID));
   mtwin=zeros(size(data.ID));
   ndtwin=ndMisodata.*(1-nottwins);
   ndnot=ndMisodata.*nottwins;
   for k=1:length(data.ID)
      lnot(k)=sum(ndnot(grainID==k)>0);
      ltwin(k)=sum(ndtwin(grainID==k)>0);
      avenot(k)=sum(ndnot(grainID==k))/lnot(k);
      avetwin(k)=sum(ndtwin(grainID==k))/ltwin(k);
      mtwin(k)=mean(ndtwin(grainID==k));
      if avenot(k)>2*avetwin(k) && mtwin(k)~=0
         disp(['grain ',num2str(k)])

         ndtemp=ndMisodata;
         ndtemp(grainID~=data.ID(k))=0;
         gtemp=zeros(size(ndtemp));
         gtemp(grainID==k)=1;
         figure(2);
         subplot(1,2,1);imagesc(ndtemp);title('Pick point for
average Euler angles (Click left of map to not change)');
         caxis([0 90]);axis equal tight
         nottwinsg=nottwins;
         nottwinsg(ndtemp==0)=0;
         subplot(1,2,2);imagesc(nottwinsg);title('red is untwinned
area');axis equal tight
         [cx ry]=ginput(1);

         close 2
         cx=round(cx);ry=round(ry);
         if cx>0 && ry>0
            nottwins(grainID==k)=1-nottwins(grainID==k);%
Switch parent and daughter
            % tpinfo 29 -parent  76 -daughter
            tpinfo(grainID==k)=nottwins(grainID==k)*29+(1-
nottwins(grainID==k))*76;
            pnew=gtemp.*nottwins==1;

[data.phis(i,1),data.phis(i,2),data.phis(i,3)]=AverageEuler(phi1(pne
w),PHI(pnew),phi2(pnew));
            changednt(k)=1;
         end
      end

   end
end
% figure;
subplot(1,2,2);imagesc(nottwins);
hold on
plot(xgbs,ygbs,'k*','MarkerSize',2)
axis equal tight
figure;imagesc(tpinfo);axis equal tight
hold on
plot(xgbs,ygbs,'k*','MarkerSize',2)
% subplot(1,2,2);imagesc(grainID);axis equal tight
% twinparent2=zeros(size(twinparent));
% twinparent2(tpinfo==29,:)=[0,0,255];
% twinparent2(tpinfo==76,:)=[255,0,0];
% twinparent2(tpinfo==255,:)=[255,255,255];

for i=1:length(data.ID)
   if changednt(i)==1


   end
end


data.rdmiso=zeros(size(data.ID));
data.tdmiso=zeros(size(data.ID));
data.ndmiso=zeros(size(data.ID));
for k=1:length(data.ID)
   rdtemp=rdMisodata;
   rdtemp(grainID~=data.ID(k))=0;
   tdtemp=tdMisodata;
   tdtemp(grainID~=data.ID(k))=0;
   ndtemp=ndMisodata;
   ndtemp(grainID~=data.ID(k))=0;

data.rdmiso(k,1)=sum(sum(rdtemp.*nottwins))/sum(sum((rdtemp.
*nottwins)>0));

data.tdmiso(k,1)=sum(sum(tdtemp.*nottwins))/sum(sum((tdtemp.*
nottwins)>0));

data.ndmiso(k,1)=sum(sum(ndtemp.*nottwins))/sum(sum((ndtemp
.*nottwins)>0));
end


for i=1:length(data.ID)
   A=euler2gmat(data.phis(i,1),data.phis(i,2),data.phis(i,3));
   nbrsizes=0;
   nbrmiso=0;
   for j=1:length(data.nbrs(1,:))
      if data.nbrs(i,j)~=0
         % Average neighbor sizes
         nind=find(data.ID==data.nbrs(i,j));
         nbrsizes=nbrsizes+data.size(nind);
         % Neighbor misorientations

B=euler2gmat(data.phis(nind,1),data.phis(nind,2),data.phis(nind,3)
);
         nbrmiso=nbrmiso+GeneralMisoCalc(A,B,lattice);
      end
   end
   data.nbrsizes(i,1)=nbrsizes/data.numnbrs(i);
   data.nbrmiso(i,1)=nbrmiso/data.numnbrs(i);

end

% Find grains that contain twins
data.twins=zeros(size(data.ID));
for i=1:length(data.ID)
   if length(unique(tpinfo(grainID==data.ID(i))))>1
      data.twins(i)=1;
   end
end
% Check to see if correct
twincheck1=zeros(size(tpinfo));
twincheck1(tpinfo==29)=1;
twincheck1(tpinfo==76)=1;
figure;imagesc((twincheck1)); shading flat
hold on
plot(xgbs,ygbs,'*k','MarkerSize',2)
```

```
twincheck2=zeros(size(tpinfo));
for i=1:numy
   for j=1:numx
      if data.twins(find(grainID(i,j)))==1
         twincheck2(i,j)=1;
      end
   end
end
% figure;pcolor(flipud(twincheck2)); shading flat

tc=twincheck2-twincheck1;
% figure; pcolor(flipud(tc)); shading flat
% Appears to be small discrepancy between 2 twin methods

% Read in Schmid and Taylor factor images for conversion
% Requires reading in parent data to exclude info from twin
% % Schmid
schmid=imread([spath sname]);
schmiddata=rgb2gray(schmid);
schmiddata=double(schmiddata);

% prompt={'Lower Limit','Upper Limit'};
% name='Input limits of Schmid factor';
% numlines=1;
% defaultanswer={'0','0.5'};
%
% slimits=inputdlg(prompt,name,numlines,defaultanswer);
smin=0;
smax=0.5;
srange=smax-smin;
schmiddata=schmiddata*smax/255;

data.schmidb=zeros(size(data.ID));
for k=1:length(data.ID)
   stemp=schmiddata;
   stemp(grainID~=k)=0;

data.schmidb(k,1)=sum(sum(stemp.*nottwins))/sum(sum((stemp.*
nottwins)>0));
end

schmid1=imread([spath1 sname1]);
schmiddata1=rgb2gray(schmid1);
schmiddata1=double(schmiddata1)*smax/255;
data.schmidca=zeros(size(data.ID));
for k=1:length(data.ID)
   stemp=schmiddata1;
   stemp(grainID~=k)=0;

data.schmidca(k,1)=sum(sum(stemp.*nottwins))/sum(sum((stemp.
*nottwins)>0));
end
figure;subplot(1,2,1);imagesc(schmiddata1);axis equal tight
title('image file')

load([spath2 sname2]);% use imread for image file
schmiddata2=sftwinmap;
if ~min(size(nottwins)==size(sftwinmap))
   disp('Schmid Factor file different size than other files')
   schmiddata2(end,:)=[];%take out last row since this is the one
missing from other data
end
% schmiddata2=rgb2gray(schmid2);
% schmiddata2=double(schmiddata2)*smax/255;
data.schmidt=zeros(size(data.ID));
for k=1:length(data.ID)
   stemp=schmiddata2;
   stemp(grainID~=k)=0;

data.schmidt(k,1)=sum(sum(stemp.*nottwins))/sum(sum((stemp.*
nottwins)~=0));
end
subplot(1,2,2);imagesc(schmiddata2);axis equal tight
title('mat file')

% schmid3=imread([spath3 sname3]);
% schmiddata3=rgb2gray(schmid3);
% schmiddata3=double(schmiddata3)*smax/255;
% data.schmidti=zeros(size(data.ID));
% for k=1:length(data.ID)
%    stemp=schmiddata1;
%    stemp(grainID~=k)=0;
%
data.schmidca(k,1)=sum(sum(stemp.*nottwins))/sum(sum((stemp.
*nottwins)>0));
% end

% Taylor
% file must be saved with range in name at set locations
% Example: NDtwinstaylor1p12345to2p12345.tif
% numbers must be in same locations
% [tyname typath]=uigetfile({'*.tif';'*.bmp'},'Select Taylor image
file',testpath);
% taylor=imread([typath tyname]);
% taylordata=rgb2gray(taylor);
% taylordata=double(taylordata);
%
% tmintemp=[tyname(14),'.',tyname(16:20)];
% tmaxtemp=[tyname(23),'.',tyname(25:29)];
% tmin=str2double(tmintemp);
% tmax=str2double(tmaxtemp);
% trange=tmax-tmin;
%
% taylordata=(taylordata*trange/255)+tmin;
% taylordata(nottwins==0)=0;
% % figure;pcolor(flipud(taylordata)); shading flat; colorbar
%
% data.taylorgrain=zeros(size(data.ID));
% for k=1:length(data.ID)
%
data.taylorgrain(k,1)=sum(taylordata(grainID==k))/sum(nottwins(
grainID==k));
% end

% Calculate self misorientation

Miso=imread([mypath myname]);
Misodata=rgb2gray(Miso);
Misodata=double(Misodata);

% mmaxtemp=[myname(12:13),'.',myname(15:18)];
mmax=5;
mrange=mmax;

Misodata=(Misodata*mrange/255);
% figure;pcolor(flipud(Misodata)); shading flat; colorbar

data.Misograin=zeros(size(data.ID));
for k=1:length(data.ID)
   mtemp=Misodata;
   mtemp(grainID~=k)=0;

data.Misograin(k,1)=sum(sum(mtemp.*nottwins))/sum(sum((mte
mp.*nottwins)>0));

   data.twinfraction(k,1)=1-
(sum(sum(nottwins(grainID==k)))/sum(sum(grainID==k)));
end
```

```
if DD
   load([ddpath ddname]);
   dd=reshape(alpha_data.alpha_total9,[Settings.data.cols
Settings.data.rows])';
   for k=1:length(data.ID)
      dtemp=dd;
      if ~min(size(nottwins)==size(dtemp))
         dtemp(end,:)=[];%take out last row since this is the one
missing from other data
      end
      dtemp(grainID~=k)=0;

data.Disloc(k,1)=sum(sum(dtemp.*nottwins))/sum(sum((dtemp.*n
ottwins)>0));
      if isnan(data.Disloc(k,1))
         data.Disloc(k,1)=1;% equals 0 after log10
      end
   end

else
   dd=ones(size(grainID)); % default for no data
   for k=1:length(data.ID)
      dtemp=dd;
      dtemp(grainID~=k)=0;

data.Disloc(k,1)=sum(sum(dtemp.*nottwins))/sum(sum((dtemp.*n
ottwins)>0));

   end
end
data.Disloc=log10(data.Disloc);

end

%% Find Boundary points for individual boundaries (including
boundary region)
% Information about boundaries stored in gbts and gbptsthin

gbiso=zeros(size(grainID));
reconatps=gbiso;
recontwins=gbiso;
g1taylor=zeros(size(data.nbrs));
g2taylor=zeros(size(data.nbrs));
% Insert Schmid data
g1schmidb=zeros(size(data.nbrs));
g2schmidb=zeros(size(data.nbrs));
g1schmidca=zeros(size(data.nbrs));
g2schmidca=zeros(size(data.nbrs));
g1schmidt=zeros(size(data.nbrs));
g2schmidt=zeros(size(data.nbrs));

g1dd=zeros(size(data.nbrs));
g2dd=zeros(size(data.nbrs));
g1miso=zeros(size(data.nbrs));
g2miso=zeros(size(data.nbrs));
gblength=zeros(size(data.nbrs));
straightness=zeros(size(data.nbrs));
RDangle=straightness;
gbtwins=zeros(size(data.nbrs));
dtaylor=zeros(size(data.nbrs));
dschmidb=zeros(size(data.nbrs));
dschmidca=zeros(size(data.nbrs));
dschmidt=zeros(size(data.nbrs));

ddd=zeros(size(data.nbrs));
rdd=zeros(size(data.nbrs));
atp=zeros(size(data.nbrs));
gbtwin1=zeros(size(data.nbrs));
gbtwin2=zeros(size(data.nbrs));
gbangle=zeros(size(data.nbrs));
gbregsize=zeros(size(data.nbrs));
edges=zeros(size(data.nbrs));

nd1=zeros(size(data.nbrs));
nd2=zeros(size(data.nbrs));

dSE=strel('diamond',1);
allgbs=zeros(size(nottwins));
gbrpred=allgbs;

idinfo=regionprops(grainID,'BoundingBox','Orientation','PixelList'
,'Perimeter','Image');

for i=1:length(data.ID)
   g1=zeros(size(grainID));
   g1(grainID==data.ID(i))=1;

   for j=1:length(data.nbrs(1,:))
      if data.nbrs(i,j)~=0

         if data.edge(i)==1 &&
data.edge(data.ID==data.nbrs(i,j))==1
            edges(i,j)=1;
         end

      % Calculate grain boundary length (straight and actual)
%       gcomb=zeros(size(grainID));

      g2=zeros(size(grainID));
      g2(grainID==data.nbrs(i,j))=1;
      gcomb=g1+g2;
      pcomb=regionprops(gcomb,'Perimeter');
      % Actual grain boundary length in um

gblength(i,j)=(idinfo(data.ID(i)).Perimeter+idinfo(data.nbrs(i,j)).Pe
rimeter-pcomb.Perimeter)/2*step+2*step;
      % Straightness % R squared value when fit to a straight line
%       gb1=zeros(size(grainID));
%       gb2=gb1;
      gb2=bwperim(g2);
      gb2d=imdilate(gb2,dSE);
      gb1=bwperim(g1);
      gb1d=imdilate(gb1,dSE);

      gb=gb1d.*gb2d.*gcomb;
      gbskel=gb;



      [gx gy]=find(gb==1);
      if length(gx)<2
%          disp('too few points to fit line')
%          disp(strcat('i = ',num2str(i),'; j = ',num2str(j)))
         length(gx);
         straightness(i,j)=0;% Indicates point boundary
         RDangle(i,j)=-1;% Indicates point boundary+
      else
         coefs=polyfit(gx,gy,1);
         gycalc=coefs(1)*gx+coefs(2);
         % straightness measure R^2
         straightness(i,j)=sum((gycalc-mean(gy)).^2)/sum((gy-
mean(gy)).^2);
         RDangle(i,j)=acos(1/sqrt(1+coefs(1)^2))*180/pi;
         % RDangle(i,j)=atan(coefs(1))*180/pi;
         if isinf(straightness(i,j))
            straightness(i,j)=0;
            RDangle(i,j)=-1;% Indicates point boundary+
            disp(['grain boundary between ',num2str(data.ID(i)),' and
',num2str(data.nbrs(i,j))])
```

```
        end
        % angle between straight boundary and x

    end

    % Define boundary regions on both sides of the boundary
    regsize=round(2/step);% approximately a 2 um region
distance from boundary
    if regsize<2
        regsize=2;
    end
    SE=strel('disk',regsize,0);
    gbregion=imdilate(gbskel,SE);
    gbregion=gbregion.*(g1+g2);
    g1reg=gbregion.*g1;
    g2reg=gbregion.*g2;
    gbregsize(i,j)=sum(gbregion(:));

    % Find properties in the grain boundary regions
    % Taylor factor, Schmid factor, Local misorientation,
dislocation
    % density
%       g1taylor(i,j)=mean(taylordata(g1reg==1));
%       g2taylor(i,j)=mean(taylordata(g2reg==1));
    % Insert Schmid data


    g1dd(i,j)=sum(dd(g1reg==1))/gblength(i,j);
    g2dd(i,j)=sum(dd(g2reg==1))/gblength(i,j);
    g1miso(i,j)=mean(Misodata(g1reg==1));
    g2miso(i,j)=mean(Misodata(g2reg==1));

%       dtaylor(i,j)=abs(g1taylor(i,j)-g2taylor(i,j));
    % Insert Schmid data


    ddd(i,j)=abs(g1dd(i,j)-g2dd(i,j));
    rdd(i,j)=max(g1dd(i,j),g2dd(i,j))/min(g1dd(i,j),g2dd(i,j));


    allgbs=allgbs+gb.*double(tpinfo);
    g1reg=gb.*g1;
    g2reg=gb.*g2;
    numtypes=unique((double(gb).*double(tpinfo))==76);
%       figure(11);imagesc(double(gbregion).*double(tpinfo));axis
equal tight
%       figure(12);imagesc(allgbs);axis equal tight


    if max(numtypes)==1 %
%           disp(['Twin touching grain: ',num2str(data.ID(i)),' and
',num2str(data.nbrs(i,j)),' boundary']);
        gbtwins(i,j)=1;
        gb1twin=max(max((double(g1reg).*double(tpinfo))==76));
        gb2twin=max(max((double(g2reg).*double(tpinfo))==76));

        if gb1twin==1 % identify atp or which grain has twin
            if gb2twin==1
                atp(i,j)=1;
                gbtwin2(i,j)=1;
            end
            gbtwin1(i,j)=1;
        else
            gbtwin2(i,j)=1;
        end
    end
    clear numtypes
    recontwins(gb==1)=gbtwins(i,j);
    reconatps(gb==1)=atp(i,j);
%       twinends=nottwins;
%       twinends((recontwins-reconatps)>0)=0.5;
%       figure(10);imagesc(twinends)
%       title('Boundaries with Twin Ends')
%
%       hold on
%       plot(xgbs,ygbs,'*k','MarkerSize',2)
%       axis equal tight
%
%       close 10
    gbangle(i,j)=idinfo(data.ID(i)).Orientation;

    % Calculations away from boundary
    % These are done to avoid noise at the boundary

    g1schmidb(i,j)=mean(schmiddata((g1.*nottwins)==1));
    g2schmidb(i,j)=mean(schmiddata((g2.*nottwins)==1));
    g1schmidca(i,j)=mean(schmiddata1((g1.*nottwins)==1));
    g2schmidca(i,j)=mean(schmiddata1((g2.*nottwins)==1));
    g1schmidt(i,j)=mean(schmiddata2((g1.*nottwins)==1));
    g2schmidt(i,j)=mean(schmiddata2((g2.*nottwins)==1));

    dschmidb(i,j)=abs(g1schmidb(i,j)-g2schmidb(i,j));
    dschmidca(i,j)=abs(g1schmidca(i,j)-g2schmidca(i,j));
    dschmidt(i,j)=abs(g1schmidt(i,j)-g2schmidt(i,j));

    A=euler2gmat(data.phis(i,1),data.phis(i,2),data.phis(i,3));
    nbrLoc=find(data.ID==data.nbrs(i,j));

B=euler2gmat(data.phis(nbrLoc,1),data.phis(nbrLoc,2),data.phis(n
brLoc,3));
    gbmiso(i,j)=GeneralMisoCalc(A,B,lattice);

    nd1(i,j)=data.ndmiso(i);
    nd2(i,j)=data.ndmiso(nbrLoc);

    % GB local region data
    gbexact=bwmorph(gb,'skel','Inf');
    gbclean=bwmorph(gbexact,'spur',2);

%       if i==37 && j==10
%           keyboard
%       end

    [order gbrough]=PntSort(gbclean);
    if gbrough==-1
        disp([data.ID(i),data.nbrs(i,j)])
    end
    gbpos=1:regsize*sqrt(3):size(order,1);
    gbpos=round(gbpos);
    if gbpos(end)~=size(order,1)
        gbpos(end+1)=size(order,1);
    end
    gbcens=order(gbpos,:);
    numsegs=size(gbcens,1);
    randreg=ceil(rand(1)*numsegs);
%       for ns=1:numsegs
%           gbr=zeros(size(gb));
%           gbr(gbcens(ns,2),gbcens(ns,1))=1;
%           gbr=imdilate(gbr,SE);
%           gbrtwin(i,j,ns)=sum((double(tpinfo(:)).*gbr(:))==76)>0;
%           % grain 1
%
gbr_bsf(i,j,ns)=mean(schmiddata((gbr.*g1.*nottwins)==1));
%
gbr_casf(i,j,ns)=mean(schmiddata1((gbr.*g1.*nottwins)==1));
%
gbr_tsf(i,j,ns)=mean(schmiddata2((gbr.*g1.*nottwins)==1));
% %           gbr_dd
% %           gbr_miso
```

```
% %           [g1phi1,g1PHI,g1phi2]=AverageEuler(phi1,PHI,phi2)
%           % grain 2
%
%           if sum(unique(tpinfo(gbr==1))==76)==1
% %            if region meets tree requirement
% %                 plot so that all predicted regions are visible
%           gbrpred=gbrpred+gbr;
%           gbrpred(gbrpred>1)=1;
%            gbri=gbrpred;
%            gbri(gbrpred==0)=.2;
%            figure1=figure(22);
%            axes1=axes('YDir','reverse','Layer','top','Color',[.2 .2
.2]);
%            hold(axes1,'all');
%            imagesc(schmiddata,'AlphaData',gbri);
%            axis equal tight
%            hold on;plot(xgbs,ygbs,'k.')
%           set(gca, 'XTick', [],'YTick',[]);
%           drawnow
%           end
%       end

      end
   end

end
figure;imagesc(grainID);
hold on
plot(xgbs,ygbs,'*k','MarkerSize',2)
axis equal tight

rttemp=gbpts*.5;
rttemp(recontwins==1)=1;
ratemp=gbpts*.5;
ratemp(reconatps==1)=1;
figure;imagesc(rttemp)
title('Boundaries with Twins')
figure;imagesc(ratemp)
title('Boundaries with ATPs')

twinends=tpinfo;
twinends((recontwins-reconatps)>0)=150;
figure;imagesc(twinends)
title('Boundaries with Twin Ends')

hold on
plot(xgbs,ygbs,'*k','MarkerSize',2)
axis equal tight
%% Organize the data into an exportable format for machine
learning

nbrtemp=data.nbrs';
nbrtemp=nbrtemp(:);
g1misotemp=g1miso';
g1misotemp=g1misotemp(:);
g2misotemp=g2miso';
g2misotemp=g2misotemp(:);
% g1taylortemp=g1taylor';
% g1taylortemp=g1taylortemp(:);
% g2taylortemp=g2taylor';
% g2taylortemp=g2taylortemp(:);
% dtaylortemp=dtaylor';
% dtaylortemp=dtaylortemp(:);
% Insert Schmid data
g1schmidbtemp=g1schmidb';
g1schmidbtemp=g1schmidbtemp(:);
g2schmidbtemp=g2schmidb';
g2schmidbtemp=g2schmidbtemp(:);
dschmidbtemp=dschmidb';
dschmidbtemp=dschmidbtemp(:);
g1schmidcatemp=g1schmidca';
g1schmidcatemp=g1schmidcatemp(:);
g2schmidcatemp=g2schmidca';
g2schmidcatemp=g2schmidcatemp(:);
dschmidcatemp=dschmidca';
dschmidcatemp=dschmidcatemp(:);

g1schmidttemp=g1schmidt';
g1schmidttemp=g1schmidttemp(:);
g2schmidttemp=g2schmidt';
g2schmidttemp=g2schmidttemp(:);
dschmidttemp=dschmidt';
dschmidttemp=dschmidttemp(:);

dd1temp=g1dd';
dd1temp=dd1temp(:);
dd2temp=g2dd';
dd2temp=dd2temp(:);

gbatemp=gbangle';
gbatemp=gbatemp(:);
straigthnesstemp=straightness';
straigthnesstemp=straigthnesstemp(:);
RDangletemp=RDangle';
RDangletemp=RDangletemp(:);
gblengthtemp=gblength';
gblengthtemp=gblengthtemp(:);
gbregtemp=gbregsize';
gbregtemp=gbregtemp(:);

gbmisotemp=gbmiso';
gbmisotemp=gbmisotemp(:);
gbtwinstemp=gbtwins';
gbtwinstemp=gbtwinstemp(:);
atptemp=atp';
atptemp=atptemp(:);
gbtwin1temp=gbtwin1';
gbtwin1temp=gbtwin1temp(:);
gbtwin2temp=gbtwin2';
gbtwin2temp=gbtwin2temp(:);

gbnd1=nd1';
gbnd1=gbnd1(:);
gbnd2=nd2';
gbnd2=gbnd2(:);

gbedges=edges';
gbedges=gbedges(:);

IDtemp=ones(size(data.nbrs));
for k=1:length(data.nbrs(1,:))
   IDtemp(:,k)=IDtemp(:,k).*data.ID;
end

IDtemp=IDtemp';
IDtemp=IDtemp(:);
gbmlinput(:,1)=IDtemp;
gbmlinput(:,2)=nbrtemp;

labels{1}='ID';
labels{2}='nbrID';

% Order of information
% grain 1, grain 2, gbmiso, gbangle, g1miso, g2miso, g1taylor,
g2taylor,
%==========
% ====
% Pick out max min ratio(max/min) and diff(max-min)
for g=1:7
   switch g
```

```
        case 1
            temp1=g1misotemp;
            temp2=g2misotemp;
            att='miso';
        case 2
            temp1=g1schmidbtemp;
            temp2=g2schmidbtemp;
            att='basalSF';
        case 3
            temp1=g1schmidcatemp;
            temp2=g2schmidcatemp;
            att='caSF';
        case 4
            temp1=g1schmidttemp;
            temp2=g2schmidttemp;
            att='twinSF';
        case 5
            temp1=g1schmidttemp./g1schmidcatemp;
            temp2=g2schmidttemp./g2schmidcatemp;
            att='catwinSF';
        case 6
            temp1=dd1temp;
            temp2=dd2temp;
            att='DisloDens';
        case 7
            temp1=gbnd1;
            temp2=gbnd2;
            att='NDmiso';
    end

    ind=(g-1)*4+3;

    gbmlinput(:,ind)=max(temp1,temp2);
    gbmlinput(:,ind+1)=min(temp1,temp2);
    gbmlinput(:,ind+2)=gbmlinput(:,ind)./gbmlinput(:,ind+1);
    gbmlinput(:,ind+3)=gbmlinput(:,ind)-gbmlinput(:,ind+1);

    labels{ind}=['max',att];
    labels{ind+1}=['min',att];
    labels{ind+2}=['ratio',att];
    labels{ind+3}=['diff',att];

end

reind=ind+4;

gbmlinput(:,reind)=straigthnesstemp;
gbmlinput(:,reind+1)=RDangletemp;
gbmlinput(:,reind+2)=gblengthtemp;
gbmlinput(:,reind+3)=gbmisotemp;

labels{reind}='straight';
labels{reind+1}='RDangle';
labels{reind+2}='length';
labels{reind+3}='GBmiso';

gbmlinput(:,reind+4)=gbtwinstemp;
gbmlinput(:,reind+5)=atptemp;
gbmlinput(:,reind+6)=gbtwin1temp;
gbmlinput(:,reind+7)=gbtwin2temp;
gbmlinput(:,reind+8)=-1*gbedges;

labels{reind+4}='twin';
labels{reind+5}='atp';
labels{reind+6}='g1twin';
labels{reind+7}='g2twin';
labels{reind+8}='edge';

gbmlinput2=sortrows(gbmlinput,2);% remove neghborless data
gbmlinput2(1:sum(gbmlinput2(:,2)==0),:)=[];
gbmlinput2=sortrows(gbmlinput2,reind+8);% remove edge data
gbmlinput2(1:sum(gbmlinput2(:,reind+8)==-1),:)=[];

rowcell=ones(length(gbmlinput2(:,1)+1),1);
colcell=ones(length(gbmlinput2(1,:)),1);

gbcell=mat2cell(gbmlinput2,rowcell,colcell);
% keyboard
% Save new file or append to already existent file
treetest=0;
qstring = 'Create new file OR append to existing file?';
choice = questdlg(qstring,'Next Action','New File','Append to
File','Test on Tree','New File');
switch choice
    case 'New File'
        % Save new file
        [filename, pathname] = uiputfile('*.xls','Save data as a Excel
file',['C:\Users\Travis\Desktop\MLtwinsNickel\' testname(1:4)]);
        xlswrite([pathname filename],[labels;gbcell]);
    case 'Append to File'
        % Append to old file
        [filename, pathname] = uigetfile('*.xls','Append data to a
previous csv file','F:\RolledNickel\');
        xlssize=size(xlsread([pathname filename],1));
        endcol=num2letter(xlssize(2));
        xlswrite([pathname
filename],gbcell,['A',num2str(xlssize(1)+1),':',endcol,num2str(xlssi
ze(1)+size(gbcell,1))]);
    case 'Test on Tree'
        treetest=1;
end

if treetest==1

    GBmiso=gbmlinput2(:,strcmp(labels,'GBmiso'));
    RDangle=gbmlinput2(:,strcmp(labels,'RDangle'));
    straight=gbmlinput2(:,strcmp(labels,'straight'));
    length=gbmlinput2(:,strcmp(labels,'length'));

    diffcatwinSF=gbmlinput2(:,strcmp(labels,'diffcatwinSF'));
    mincatwinSF=gbmlinput2(:,strcmp(labels,'mincatwinSF'));

    diffmiso=gbmlinput2(:,strcmp(labels,'diffmiso'));

    maxNDmiso=gbmlinput2(:,strcmp(labels,'maxNDmiso'));

    diffbasalSF=gbmlinput2(:,strcmp(labels,'diffbasalSF'));
    minbasalSF=gbmlinput2(:,strcmp(labels,'minbasalSF'));
    maxbasalSF=gbmlinput2(:,strcmp(labels,'maxbasalSF'));

    cnt1=0;
    cnt2=0;
    cnt3=0;
    cnt4=0;
    cnt5=0;
    cnt6=0;
    cnt7=0;
    cnt8=0;
    cnt9=0;
    cnt10=0;
    cnt11=0;
    cnt12=0;
    ptwin=0*GBmiso;

for i=1:size(GBmiso,1)
    if length(i) <= 10.052
        if GBmiso(i) <= 27.146
            if mincatwinSF(i) <= 0.738
                if diffcatwinSF(i) <= 0.317
                    if minbasalSF(i) <= 0.093
```

```
                ptwin(i)=0.5;                disp(1)
cnt1=cnt1+1;
            else
                ptwin(i)=1;                disp(2)
cnt2=cnt2+1;
            end
        else
            ptwin(i)=0.5;                disp(3)
cnt3=cnt3+1;
        end
    else
        ptwin(i)=0.5;                disp(4)
cnt4=cnt4+1;
    end
else
    if GBmiso(i) <= 46.55
        ptwin(i)=0.5;                disp(5)
cnt5=cnt5+1;
    else
        if diffbasalSF(i) <= 0.06
            if diffmiso(i) <= 0.032
                ptwin(i)=1;                disp(6)
cnt6=cnt6+1;
            else
                if RDangle(i) <= 46.63
                    ptwin(i)=0.5;                disp(7)
cnt7=cnt7+1;
                else
                    ptwin(i)=1;                disp(8)
cnt8=cnt8+1;
                end
            end
        else
            if GBmiso(i) <= 49.314
                ptwin(i)=1;                disp(9)
cnt9=cnt9+1;
            else
                ptwin(i)=0.5;                disp(10)
cnt10=cnt10+1;
            end
        end
    end
end
else
    if maxbasalSF(i) <= 0.359
        ptwin(i)=1;                disp(11)
cnt11=cnt11+1;
    else
        ptwin(i)=0.5;                disp(12)
cnt12=cnt12+1;
    end
end
end

%% Plot GB predictions
predmap=zeros(size(ratemp));
usemap=predmap;

for i=1:size(gbmlinput2(:,1),1)
    g1=zeros(size(grainID));
    g1(grainID==gbmlinput2(i,1))=1;
    g2=zeros(size(grainID));
    g2(grainID==gbmlinput2(i,2))=1;
    gcomb=g1+g2;
    gb2=bwperim(g2);
    gb2d=imdilate(gb2,dSE);
    gb1=bwperim(g1);
    gb1d=imdilate(gb1,dSE);
    gb=gb1d.*gb2d.*gcomb;
    predmap(gb==1)=ptwin(i);
    usemap(rttemp==1 & gb==1)=1;
end
figure;
imagesc(usemap)
axis equal tight
matchmap=(ratemp.*gbpts-predmap.*gbpts).*usemap;
matchmap(matchmap==0.5)=1;
matchmap(matchmap==-0.5)=0.5;
figure;imagesc(matchmap);axis equal tight
mmtemp=bwmorph(ceil(matchmap),'clean');
% figure;imagesc(mmtemp);axis equal tight

xsize=size(matchmap,2);
ysize=size(matchmap,1);
[X,Y]=meshgrid(1:xsize,1:ysize);
predmap=predmap.*usemap;
 mm2=bwmorph(mmtemp,'dilate',3);
 mm2(predmap.*usemap>0)=0;
predmap(mm2==1)=0.7;
x_miss=X(matchmap>0);
y_miss=Y(matchmap>0);
figure;
imagesc(predmap);
axis equal tight
hold on
plot(xgbs,ygbs,'k.','MarkerSize',1)

% make rgb error map
rgberror=zeros(ysize,xsize,3);
r=zeros(ysize,xsize);
g=r;
b=r;
r(predmap==0.5 | predmap==1)=1;
g(predmap==0.5 | predmap==0.7)=1;
b(predmap==0.5 | predmap==0)=1;
rgberror(:,:,1)=r;
rgberror(:,:,2)=g;
rgberror(:,:,3)=b;
figure;
imagesc(rgberror)
axis equal tight
hold on
plot(xgbs,ygbs,'k.','MarkerSize',1)
%% Choose to analyze another file or not
qstring = 'Analyze another file OR finish?';
choice = questdlg(qstring,'Another another file or not','Analyze a
New File','Finish','Finish');
switch choice
    case 'Analyze a New File'
        % Analyze another file
        close all
        clear all
        clc
        more=1;
    case 'Finish'
        % End file analysis
        more=0;
end

end
```