Artificial neural network modeling of flow stress response as a function of dislocation microstructures

Osama Yousef AbuOmar

Follow this and additional works at: https://scholarsjunction.msstate.edu/td

Recommended Citation
https://scholarsjunction.msstate.edu/td/730

This Graduate Thesis is brought to you for free and open access by the Theses and Dissertations at Scholars Junction. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholars Junction. For more information, please contact scholcomm@msstate.libanswers.com.
ARTIFICIAL NEURAL NETWORK MODELING OF FLOW STRESS RESPONSE AS A FUNCTION OF DISLOCATION MICROSTRUCTURES

By

Osama Yousef AbuOmar

A Thesis
Submitted to the Faculty of Mississippi State University in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Engineering in the Department of Electrical and Computer Engineering

Mississippi State, Mississippi

August 2007
Copyright by
Osama Yousef AbuOmar
2007
ARTIFICIAL NEURAL NETWORK MODELING OF FLOW STRESS RESPONSE AS A FUNCTION OF DISLOCATION MICROSTRUCTURES

By
Osama Yousef AbuOmar

Approved:

<table>
<thead>
<tr>
<th>Eric A. Hansen</th>
<th>Mark Horstemeyer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Associate Professor of Computer Science and Engineering (Director of Thesis)</td>
<td>Professor of Mechanical Engineering (Co-Major Professor)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Yul Chu</th>
<th>Julian Boggess</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assistant Professor of Electrical and Computer Engineering (Committee Member)</td>
<td>Associate Professor of Computer Science and Engineering (Committee Member)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nicholas H. Younan</th>
<th>Roger King</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professor of Electrical and Computer Engineering, Graduate Coordinator, Department of Electrical and Computer Engineering</td>
<td>Associate Dean for Research and Graduate Studies, Bagely College of Engineering</td>
</tr>
</tbody>
</table>
An artificial neural network (ANN) is used to model nonlinear, large deformation plastic behavior of a material. This ANN model establishes a relationship between flow stress and dislocation structure content. The density of geometrically necessary dislocations (GNDs) was calculated based on analysis of local lattice curvature evolution. The model includes essential statistical measures extracted from the distributions of dislocation microstructures, including substructure cell size, wall thickness, and GND density as the input variables to the ANN model. The model was able to successfully predict the flow stress of aluminum alloy 6022 as a function of its dislocation structure content. Furthermore, a sensitivity analysis was performed to identify the significance of individual dislocation parameters on the flow stress. The results show that an ANN model can be used to calibrate and predict inelastic material properties that are often cumbersome to model with rigorous dislocation-based plasticity models.
DEDICATION

I would like to dedicate this research to my family; my dad, mum, and brothers and to all of my sincere friends and colleagues.
ACKNOWLEDGMENTS

The author wishes to express his warm thanks to Dr. Eric Hansen for his dedicated and faithful help from the beginning to the end. Special thanks to Dr. Reza Shahbazian Yassar and Dr. Mark Horstemeyer for their continuous mentoring and valuable advices while pursuing this work. Sincere regards should be extended also to other committee members; Dr. Yul Chu and Dr. Julian Boggess for their cooperation and ultimate help during the organization phase of this work.

Sincere regards to Dr. Ricolindo Carino, Neil Williams and Kiran Solanki for their vital help in software development phase of this work.

Also, thanks to Stephen Horstemeyer for his assistance in facilitating this research at the materials characterization laboratory at the Center for Advanced Vehicular Systems (CAVS). Warm thanks also include Kyla Stolting, James C. Baird, and John Murphy from the Mechanical Engineering Department at Mississippi State University for their efforts in conducting the sample preparation and EBSD measurements. Furthermore, the author wishes to express his acknowledgments to Dr. Pankaj Trivedi from the Institute of Shock Physics at Washington State University for his assistance on GN dislocation analysis.

Finally, warm thanks to all of my friends. I believe that without their spiritual help, this work would not have come into being.
TABLE OF CONTENTS

DEDICATION ............................................................................................................. ii
ACKNOWLEDGMENTS ........................................................................................... iii
LIST OF TABLES ....................................................................................................... vi
LIST OF FIGURES ..................................................................................................... vii

CHAPTER

I. INTRODUCTION ........................................................................................... 1
   1.1 Motivation ............................................................................................ 1
   1.2 Organization of Thesis ......................................................................... 2

II. ARTIFICIAL NEURAL NETWORKS ........................................................... 3
   2.1 Introduction .......................................................................................... 3
       2.1.1 Multilayer Feedforward Artificial Neural Network .................. 3
       2.1.2 Backpropagation Training Algorithm ...................................... 6
   2.2 Predicting Flow Stress using ANNs .................................................... 11

III. DESIGN OF NEURAL NETWORK MODEL .............................................. 16
   3.1 Deformation and Dislocation Concepts ............................................... 17
   3.2 Model Algorithms ................................................................................ 21
       3.2.1 Dislocation Cell Wall Thickness Algorithm ............................ 24
       3.2.2 Dislocation Cell Size Algorithm .............................................. 28
       3.2.3 Dislocation Cell Wall Thickness and Cell Size Analysis ...... 31
       3.2.4 Dislocation Density calculation technique............................... 41
   3.3 Statistical Measures of Inputs Distributions ........................................ 43
   3.4 Model Architecture .............................................................................. 46

IV. RESULTS OF TRAINING .............................................................................. 51
   4.1 Cross-Validation Analysis: Early Stopping method ............................ 51
   4.2 Model Prediction vs. Experimental values ......................................... 54
   4.3 Bayesian Regularization: Sensitivity Analysis .................................... 61
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Topology and training parameters for the developed neural network [10] ..... 13</td>
</tr>
<tr>
<td>3.1</td>
<td>Statistical analysis sample under “under-aged” condition; instead of having 15 datasets representing three different orientations per strain, only 5 sets seem to be reasonable after computing the average value for all statistical measures under the same strain. This is applicable to “over-aged” and “peak-aged” cases as well............................................... 49</td>
</tr>
<tr>
<td>4.1</td>
<td>The predicted Flow Stress (in MPa) for each validation example used in ANN model................................................................................................ 56</td>
</tr>
<tr>
<td>4.2</td>
<td>Percentage errors for the three different cases that the model analyzed........ 59</td>
</tr>
<tr>
<td>4.3</td>
<td>The average weight of each input of our model. This indicates the level that each input affects the predicted result. Dislocation Cell Wall Thickness, Cell Size, and GND density have average weights of 2.42, 3.12, and 3.24 respectively ................................................................. 64</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Architecture of typical ANN [3]</td>
<td>5</td>
</tr>
<tr>
<td>2.2</td>
<td>Comparison of measured and calculated (ANN) stress-strain curves for different temperatures and strains [17]</td>
<td>14</td>
</tr>
<tr>
<td>3.1</td>
<td>EBSD Process; depending on the step size and the scans used in EBSD, different orientations measurements are produced [19, 21]</td>
<td>19</td>
</tr>
<tr>
<td>3.2</td>
<td>Analysis of deformed structures (misorientation boundaries); (a) Crystal orientation within different grains, (b) The boundary map represents the misorientation angle between two neighbor points greater than 1°</td>
<td>20</td>
</tr>
<tr>
<td>3.3</td>
<td>Analysis of deformed microstructures (misorientation plots)</td>
<td>20</td>
</tr>
<tr>
<td>3.4</td>
<td>Directional properties of dislocations</td>
<td>21</td>
</tr>
<tr>
<td>3.5</td>
<td>The goal is to develop a learning algorithm linking the dislocation structure parameters to the material properties</td>
<td>25</td>
</tr>
<tr>
<td>3.6</td>
<td>The trigonometric operation to compute x1 and x2 in Case 3</td>
<td>27</td>
</tr>
<tr>
<td>3.7</td>
<td>A graphical representation of Dislocation Cell Wall Thickness and Cell Size</td>
<td>30</td>
</tr>
<tr>
<td>3.8</td>
<td>Frequency count for Dislocation Cell Wall Thickness values for three different misorientation angles under “under-aged” case</td>
<td>32</td>
</tr>
<tr>
<td>3.9</td>
<td>Frequency count for Dislocation Cell Wall Thickness values for three different misorientation angles under “peak-aged” case</td>
<td>32</td>
</tr>
<tr>
<td>3.10</td>
<td>Frequency count for Dislocation Cell Wall Thickness values for three different misorientation angles under “over-aged” case</td>
<td>33</td>
</tr>
<tr>
<td>3.11</td>
<td>Frequency count for Dislocation Cell size values for three different misorientation angles under “under-aged” case</td>
<td>33</td>
</tr>
</tbody>
</table>
3.12 Frequency count for Dislocation Cell Size values for three different misorientation angles under “peak-aged” case ......................................................... 34
3.13 Frequency count for Dislocation Cell Size values for three different misorientation angles under “over-aged” case........................................... 34
3.14 The frequency of Dislocation Cell Wall Thickness under three hardening conditions for 0% strain................................................................. 36
3.15 The frequency of Dislocation Cell Wall Thickness under three hardening conditions for 5% strain................................................................. 36
3.16 The frequency of Dislocation Cell Wall Thickness under three hardening conditions for 10% strain................................................................. 37
3.17 The frequency of Dislocation Cell Wall Thickness under three hardening conditions for 15% strain................................................................. 37
3.18 The frequency of Dislocation Cell Wall Thickness under three hardening conditions for 20% strain................................................................. 38
3.19 The frequency of Dislocation Cell Size under three hardening conditions for 0% strain ...................................................................................... 38
3.20 The frequency of Dislocation Cell Size under three hardening conditions for 5% strain ...................................................................................... 39
3.21 The frequency of Dislocation Cell Size under three hardening conditions for 10% strain ...................................................................................... 39
3.22 The frequency of Dislocation Cell Size under three hardening conditions for 15% strain ...................................................................................... 40
3.23 The frequency of Dislocation Cell Size under three hardening conditions for 20% strain ...................................................................................... 40
3.24 The Process of dividing the original GND density matrix into subcells and constructing a new GND density matrix based on the average values of each subcell................................................................. 43
3.25 Model architecture (9 inputs, one hidden layer of 6 neurons, and an output layer with one output. Learning Rate = 0.15). Each quantity of dislocation structure can be represented by 3 statistical measures of Mean, Standard Deviation, and Skewness................................................. 50

4.1 Cross-Validation technique: the weights that result in the least MSE on the validation set are saved, so we can return to them after overtraining is detected ...................................................................................................... 53

4.2 Cross-Validation analysis using a validation set and a training set.............. 54

4.3 The Flow Stress versus deformation percentages or strains (0 to 20%) for the model prediction and experimental data (for peak-aged, under-aged, and over-aged samples)................................. 57

4.4 MSE versus number of epochs (running time) needed to achieve the optimum performance........................................................................................................ 58

4.5 MSE versus the number of neurons in the hidden layer needed to achieve the optimum performance.......................................................... 60

4.6 ANN model after applying Bayesian Regularization: out of 60 weights used, only 15 are effective; 9 from input to hidden layer (one per input) and 6 from hidden to output layer............................................... 62

4.7 The prediction of Flow Stress values using multiple linear regression technique under the three considered hardening conditions in all strains. The points in the Figure (not filled) demonstrate the experimental values of Flow Stress obtained in each plastic strain percentage.......................... 68
CHAPTER I
INTRODUCTION

1.1 Motivation

Modeling relationships between the observed properties of a material and its microstructure is a key step in the materials design process. But often, these relationships are too complex or not well-enough understood to model analytically. This motivates the use of machine learning techniques, including artificial neural networks, to build useful predictive models.

This thesis describes how an artificial neural network (ANN) can be used to predict the flow stress response of a precipitation-hardening aluminum alloy based on its dislocation structure content. The dislocation structure is characterized using a high-resolution electron backscatter diffraction technique. Measurements are summarized by statistical parameters of the distribution of several key features of the dislocation microstructures, including substructure cell size, wall thickness, and density of geometrically necessary dislocations. A multilayer feedforward network is trained by the backpropagation algorithm to predict flow stress response as a function of these input parameters. Artificial neural networks perform particularly well on this task because the relationship between the parameters of the dislocation structures and flow stress response is nonlinear. Results presented in this thesis show that an artificial neural network
predicts the flow stress response much more accurately than multiple linear regression.

Artificial neural networks have been successfully applied in materials science by several previous researchers. However this work appears unique in that it is the first to use dislocation density as an input microstructure variable. The results show how the materials engineering and computer science communities can collaborate by using machine learning techniques to model challenging, real-world material design problems.

1.2 Organization of Thesis

The rest of the thesis is organized as follows. Chapter 2 provides a background review of neural networks and their training algorithms, especially the backpropagation training algorithm for multilayer feedforward neural networks. It also reviews previous applications of ANNs in materials science. Chapter 3 describes the design of a neural network that predicts flow response, including a description of all dislocation microstructures algorithms that were developed, the statistical measures that were extracted from various microstructures’ distributions, and the network architecture. Chapter 4 describes the results of training the network, including three different types of analysis; sensitivity analysis, cross-validation analysis, and linear regression analysis. Chapter 5 concludes the thesis and describes potential future work.
CHAPTER II
ARTIFICIAL NEURAL NETWORKS

2.1 Introduction

Artificial neural networks are mathematical models and algorithms that emulate certain aspects of the information-processing and knowledge-gathering methods of the human nervous system [2]. The use of ANNs to perform computations is inspired by neurobiological systems in which the messages are transferred between several neurons using specific neurotransmitters. ANNs use neurons and weights in order to perform parallel and distributed processing to find out the relations between the inputs and the desired outputs [1].

An ANN can perform highly complex mappings on nonlinearly related data by inferring subtle relationships between input and output parameters. It can, in principle, generalize from a limited quantity of training data to overall trends in functional relationships.

2.1.1 Multilayer Feedforward Artificial Neural Network

Although several network architectures and training algorithms are available, the feed-forward neural network (FFNN) trained with the back-propagation (BP) learning algorithm is the most commonly used. A FFNN network is trained using a supervised
learning mechanism. In supervised learning, the network is presented with patterns, (inputs and target outputs), in the training phase. After each presentation, the weights of the network are adjusted to decrease the error between the network’s output and the target output [2].

Network computation is performed by a dense mesh of computing nodes and connections. They operate collectively and simultaneously on most or all data and inputs. The basic processing elements of neural networks are called artificial neurons, or simply neurons. Often they are simply called nodes. Neurons perform as summing and non-linear mapping junctions. In some cases, they can be considered as threshold units that fire when their total input exceeds certain bias levels. Neurons usually operate in parallel and are configured in regular architectures. They are often organized in layers, and feedback connections both within the layer and toward adjacent layers are allowed. Connection strength is expressed by a numerical value called a weight, which is modified during training [3].

Among various artificial neural networks, the elementary multilayer perceptrons (MLP) with sigmoidal transfer function have been successfully applied to solve some difficult and diverse problems [4, 5] such as non-linear discriminant function classification. The feedforward network learns from the input data by the supervision of the output data creating single linear discriminant functions by each sigmoid hidden unit and combines them. Thus, this piecewise linear discriminant function works as a non-linear discriminator.
Figure 2.1 shows an example of a multilayer feedforward neural network [3].

In conclusion, in the implementation of ANN, data are categorized as input patterns and target patterns. The input patterns are fed to the network, which then performs feed-forward computations to calculate output patterns. The output patterns are compared with corresponding target patterns and the summation of the square of the error is calculated. The error is then back-propagated through the network using the gradient-descent rule to modify the weights and minimize the summed squared error. The back-
propagation algorithm is the standard method of choice for finding the weights of feed-forward neural networks [6, 7].

As a whole, the following factors play significant roles in designing any FFNN:

- Number of hidden layers
- Number of neurons (nodes) operated in each layer
- Mean Square Error (MSE)
- Number of iterations of the program.
- Learning and Momentum rates ($\alpha$ and $\epsilon$ or $\mu$ respectively)
- Activation functions used (for input-hidden layers and for hidden-output layers)

2.1.2 Backpropagation Training Algorithm

Training the network in a supervised manner with a popular algorithm known as the error backpropagation (BP) is frequently done. BP is an optimization technique for implementing gradient descent in weight space for multilayer feedforward networks.

The basic idea of the technique is to efficiently compute partial derivatives of an approximating function $F(w;x)$ realized by the network with respect to all the elements of the adjustable weight vector $w$ for a given value of input vector $x$ and output vector $y$. The weights are adjusted to fit linear piecewise discriminant functions in feature space for the best class separability. The difference between the network output and the supervisor output is minimized according to a predefined error function (performance criterion) such as mean square error (MSE).
The general aim in the training process is to teach the relations between input and output values to the program and get the results with the lowest possible errors. In neural network applications, output values are reduced to values between 0 and 1, which is called the normalization process. This was carried out by dividing the input and output values by some real numbers. On the other hand, the training process is always performed by the ‘trial and error method’ and there is no automatic way for that when using artificial neural networks. Training iterations are made by changing the learning rates ($\alpha$), momentum values ($\varepsilon$) and number of nodes in the hidden layers. The errors are minimized by iterating the neural network programs many times by using the appropriate $\alpha$, $\varepsilon$ and hidden layer nodes. So, training iterations take time because of the many trials undertaken by changing the parameters of programs ($\alpha$, $\varepsilon$ and hidden layer node numbers). Training is finished when the optimum parameters (like parameters that result in least MSE) are determined [3].

Back-propagation training algorithm is an iterative gradient technique that minimizes a cost function equal to the mean square difference between the desired and the actual net outputs. According to [10], the design procedure of the neural network covers the following steps:

Consider a network that consists of input vectors $x_0$, $x_1$, \ldots $x_N$ and the desired output vectors $y_0$, $y_1$, \ldots $y_M$. The output of each node $h_i$ in each hidden layer can be expressed according to the following relationship:
\[ h_i = f \left( \sum_{j=1}^{N} w_{ij} x_j - \vartheta_i \right) \]  

(2.1)

where \( w_{ij} \) are the weights between input and hidden layers and \( \vartheta_i \) is the bias between input and hidden layers; \( f \) is the activation (transfer) function. Each set of neurons and epochs may have a different activation function than the other set. The activation functions may be any differentiable transfer functions such as linear, hyperbolic tangent sigmoid or log sigmoid.

After multiple iterations to minimize the MSE of the network output, hyperbolic tangent sigmoid (Eq. 2.2) and log sigmoid (Eq. 2.3) transfer functions were applied to the hidden and the output layers respectively.

\[ Tansig(n) = \frac{2}{(1 + \exp(-2 \cdot \Sigma))} - 1 \]  

(2.2)

\[ Logsig(n) = \frac{1}{(1 + \exp(-\Sigma))} \]  

(2.3)

where \( \Sigma \) is the ANN model input to the hidden node in the case of the hyperbolic tangent sigmoid (Equation 2.2) and the output value of the hidden node in the case of log sigmoid (Equation 2.3).

At the beginning all weights and biases were set to random values. In the next step the output of each node in the output layer was computed according to:

\[ y_k = f \left( \sum_{j=1}^{M} w_{kj} x_j - \vartheta_k \right) \]  

(2.4)
where $y_k$ is the vector of the output layer nodes, $w_{kj}$ are the weights between hidden and output layers and $\vartheta_k$ is the bias between hidden and output layers.

The mean square error can be expressed by the relation:

$$V = \frac{1}{2} \sum_{q=1}^{Q} (t_q - y_q^L)^T (t_q - y_q^L) = \frac{1}{2} \sum_{q=1}^{Q} e_q^T e_q$$  \hspace{1cm} (2.5)$$

where $y_q^L$ is the output of the network corresponding to the $q$th input $x_Q$ at layer $L$, $t_q$ is the target, and $e_q = (t_q - y_q^L)$ is the error term.

The weight update is performed after the presentation of all training data (batch training). The weight update for the steepest descent algorithm is:

$$\Delta w_{k,i,j}^i = -\alpha \frac{\partial V}{\partial w_{k,i,j}}$$ \hspace{1cm} (2.6)$$

$$\Delta \vartheta_k^i = -\alpha \frac{\partial V}{\partial \vartheta_k^i}$$ \hspace{1cm} (2.7)$$

where $\alpha$ is the learning rate which corresponds to the rate of convergence between the current solution and the global minimum. Effects of changes in the net input of neuron $i$ in layer $k$ to the performance index are defined as the sensitivity:

$$\delta_i^k \equiv \frac{\partial V}{\partial n_i^k}$$ \hspace{1cm} (2.8)$$

The error between target and observed output can be calculated according to:

$$\delta = y_k (1 - y_k) (\hat{y}_k - y_k)$$ \hspace{1cm} (2.9)$$
where $\delta_k$ is the errors vector for each output node and $Y_k$ is the target of output layer.

Hence, the hidden layer error is:

$$\delta = h_i(1 - h_i) \sum_{k=1}^{m} \delta_k w_{kj}$$

(2.10)

where $\delta_i$ is the errors vector for each hidden layer and $m$ is the number of nodes in each hidden layer. In the next step the weights and biases in the output layer should be adjusted,

$$w_{yi}(t + 1) = w_{yi}(t) + \alpha \delta_i h_i + \mu (w_{yi}(t) - w_{yi}(t - 1))$$

(2.11)

$$\theta_i(t + 1) = \theta_i(t) + \alpha \delta_i$$

(2.12)

where $\mu$ is the momentum that helps the network overcome local minima and $t$ is an indication of the weight and bias states (i.e., $t-1$ represents the last state of weight and bias, $t$ is the current state, and $t+1$ is the next state). Following the output layer, the weights and biases in the hidden layer need to be adjusted,

$$w_{ji}(t + 1) = w_{ji}(t) + \alpha \delta h_i + \mu (w_{ji}(t) - w_{ji}(t - 1))$$

(2.13)

$$\theta_j(t + 1) = \theta_j(t) + \alpha \delta$$

(2.14)

This process should be repeated until the output layer error is within the specified tolerance. For fast optimization, we used the Levenberg–Marquardt training algorithm. The training typology is the batch mode, which means that the network weights and biases are updated only after the entire training set has been applied to the network, so the predicted results are more accurate than the standard BP algorithm in which the weights and biases are updated after applying each training example to the network.
The average error between experimental and predicted results was defined as follows:

\[
Average\ error\% = \frac{\sum_{j=1}^{N}((\text{exp}_j - \text{pred}_j) / \text{exp}_j)}{N} \times 100\%
\]  

(2.15)

where \( N \) is the total number of experimental data of testing conditions.

## 2.2 Predicting Flow Stress using ANNs

There are several applications of neural networks and other artificial intelligence techniques in material science. These applications include stress-strain modeling [7], grain size discrimination applications [12], microstructure evolution and modeling using several mechanical processes [3, 13], correlating mechanical properties of several alloys with the composition and heat treatment parameters [14], prediction of alloys’ rheological behavior under certain conditions [10], prediction of grain size of several alloys [2], and analysis of the microstructure of Atmospheric Plasma Spraying (APS) coatings [15]. Furthermore, numerous examples of the application of ANNs in metal forming can be found in the scientific literature. Applications include control of rolling mills, prediction of yield strength in plate mills [8], rolling loads, plate bending in asymmetrical rolling, and roll bending in four-high stands [9, 10].

In one of the earliest attempts to apply ANNs in material science, Ellis, Yao, Zhao, and Penumadu et al. [7] implemented ANNs for modeling the stress-strain relationship of sands with varying grain size distribution and stress history. This work
demonstrated the ability of neural networks to simulate unload-reload loops of the soil stress-strain characteristics. The inputs of the neural network are the confining pressure ($\sigma_{3c}'$), initial relative density ($D_r$), and over consolidation ratio (OCR) which reflects the previous stress history. The output of the system is the grain size distribution on stress-strain relationship of Mortar Sand.

In the work of Bariani, Bruschi, and Dal Negro et al. [10], neural networks are utilized to represent the rheological behavior of the Nickel-base superalloy Nimonic 80A under deformation conditions approximating thermo-mechanical cycles of industrial hot forging operations. Instantaneous values of equivalent strain, strain and temperature have been first employed as network inputs and the effects of the previous thermal and mechanical cycles on material flow stress have been taken into account adding to the network inputs the microstructural parameter grain size that has been proven to suitably represent this influence. The equivalent stress is the only output. Table 2.1 shows the design details of this application done in [10].

Recent approach uses ANNs in a Bayesian framework [16]. This approach was used to predict flow stress from the inputs of composition, temperature, strain rate and strain. The Bayesian approach to neural networks makes predictions with error bars, with the magnitude depending upon the position in the input space and perceived level of noise in the model.

Regarding the dislocation density applications using artificial intelligence techniques, Kusiak and Pietrzyk et al. [17] have presented an approach that applies
artificial neural networks to the prediction of yield stress in hot forming of metals. The task of the network is to predict a time-derivative of the dislocation density during hot deformation. The inputs are the state of the material defined by the current dislocation density and by the time-integral of strain, the current strain rate and temperature. The flow stress curve is determined from the dislocation density vs. strain function, which is calculated using a finite difference technique in which the time-derivative of the dislocation density is supplied by the artificial neural network. The main objective was to extend the analysis to different strain rates. Training of the network was done using experimental data from results of axi-symmetrical compression tests performed at three temperatures and three strain rates. Figure 2.2 shows the analysis that resulted from the ANN model [17].

Table 2.1 Topology and training parameters for the developed neural network [10]

<table>
<thead>
<tr>
<th>Number of input nodes</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of output nodes</td>
<td>1</td>
</tr>
<tr>
<td>Number of hidden layers</td>
<td>1</td>
</tr>
<tr>
<td>Number of hidden nodes</td>
<td>7</td>
</tr>
<tr>
<td>Activation function input-hidden layers</td>
<td>Sigmoid</td>
</tr>
<tr>
<td>Activation function hidden-output layers</td>
<td>Linear</td>
</tr>
<tr>
<td>Number of epochs</td>
<td>50</td>
</tr>
<tr>
<td>Learning rate</td>
<td>0.9</td>
</tr>
<tr>
<td>( \mu_{\text{dec}} )</td>
<td>0.1</td>
</tr>
<tr>
<td>( \mu_{\text{inc}} )</td>
<td>10</td>
</tr>
<tr>
<td>( \mu_{\text{max}} )</td>
<td>1.00E+10</td>
</tr>
</tbody>
</table>
Figure 2.2 Comparison of measured and calculated (ANN) stress-strain curves for different temperatures and strains [17]

Another recent application of ANNs in modeling the constitutive flow behavior of as cast 304 stainless steel under hot deformation [11]. The inputs of the neural network are strain, strain rate, and temperature, and flow stress is the output. Experimental data obtained from hot compression tests in the temperature range 1023-1523 K, strain range
0.1-0.5, and strain rate range $10^{-3}$-10$^2$ s$^{-1}$ are employed to develop the model. An attempt is made to quantify the extrapolation ability of the developed network.

Aside from this, there are few examples of dislocation density applications using artificial intelligence techniques and applications of ANNs using dislocation cell wall thickness, dislocation cell size, and grain refinement were not found in the literature. So, further research must be done in these crucial areas.
CHAPTER III
DESIGN OF NEURAL NETWORK MODEL

This chapter describes how to use the ANN technique to model some of the microstructure properties in “deformed” structures of the material, such as metals, to predict the amount of stress or pressure (in MPa) needed to deform the material based on those microstructure properties under four different values of strains (0%, 5%, 10%, 15%, and 20%).

We have analyzed three microstructure properties in deformed structures which are:

- Dislocation Cell Wall Thickness
- Dislocation Cell Size
- Dislocation Density

Before analyzing these three properties and describing the algorithms used to compute each of them and the structure of the ANN that is used to model them and predict the required stress, let’s provide an overview of the deformation and dislocation concepts.
3.1 Deformation and Dislocation Concepts

Any material consists of a number of microstructural units called “grains”. The shapes and characteristics of these grains are obtained using a technique called electron backscatter diffraction (EBSD) [18] in which a beam of electrons is concentrated toward an area of interest in the material sample and then different crystal orientations of these grains will be obtained using a graphical display. EBSD allows rapid orientation mapping over large sample areas of the material into units called grains and these local crystallographic orientations are measured in a scanning electron microscope (SEM) equipped with an ultra-sensitive camera for recording the crystal orientation of grain patterns resulting from EBSD technique as in Figure 3.1 [21].

Now, after applying a stress (that results in certain reduction of the original size of the sample), the sample will be deformed and the dislocations (subcells) will be produced within the grains. Those dislocation structures are analyzed using a transmission electron microscopy technique (TEM) which is an imaging technique whereby a beam of electrons is focused onto a specimen causing an enlarged version to appear on a fluorescent screen or layer of photographic film, or to be detected by a CCD camera [19]. An EBSD image of the grains and the dislocation structures is illustrated in Figure 3.2.

As discussed in [18], during plastic deformation of polycrystalline materials, individual grains are sometimes subdivided into crystallites rotating independently of one another to accommodate the imposed strain. Furthermore, the reason for the grain fragmentation that we mentioned above is that the number and selection of
simultaneously acting slip systems differs among neighboring volume elements within a
grain. This leads to differences in lattice rotations between neighboring elements within a
grain when the material is strained. So, depending upon the crystal lattice orientation of
the grain and its interaction with near neighbors, grains could develop a well-defined cell-
block structure of similar orientation but rotating at different rates and sometimes in
different directions resulting in an angle called the misorientation angle, and this angle
may differ between each pair of points or precipitates in one grain or in different grains.

This may lead us to think that the relation between the distance between each pair
of points (measured in microns), within one grain or different grains, and the
misorientation angle will be the basis for all the microstructure modeling in deformed
structures. There are two different ways in the literature to express this distance either
point-point distance or point-origin distance. The latter considers one point as a reference
and measures the distances of all points up to the reference points, so the distances will
be accumulated along the grains until we finish analyzing the material sample. The
former considers the distance between each pair of points as independent values
regardless of the distances between other pairs of points. As indicated in [20], point-point
distance is the most common and convenient method in modeling microstructure
properties of the material.

For the orientation between two points in the sample to be acceptable, the
misorientation angle should be between 1° and 15° within one grain because the points
that have a misorientation angle of less than 1° may be neglected since we will not have a
good distinguished orientation of these points. Delannay, Logé, Chastel, Signorelli, and Houtte et al. [20] prove that large disorientations (>15°) build up within some grains, already after having undergone moderate rolling strains. This means that whenever we have a point in a sample of misorientation angle greater than 15°, then it will be the starting point of the next grain. For that reason, for analysis within one grain, the range of 1°-15° must be considered.

Figure 3.1 EBSD Process; depending on the step size and the scans used in EBSD, different orientations measurements are produced [19, 21]

Figure 3.3 is a graphical representation of point-point distance versus the misorientation angle of a probe that passes through one grain. The misorientation plot shows the misorientation angle between two neighbor points along the probe line.
Figure 3.2 Analysis of deformed structures (misorientation boundaries); (a) Crystal orientation within different grains, (b) The boundary map represents the misorientation angle between two neighbor points greater than 1°.

Figure 3.3 Analysis of deformed microstructures (misorientation plots)
As we mentioned above, individual grains are sometimes subdivided into crystallites rotating independently of one another. This can be shown experimentally in Figure 3.4.

![Figure 3.4 Directional properties of dislocations](image)

**Figure 3.4 Directional properties of dislocations**

### 3.2 Model Algorithms

As mentioned above, we are mainly interested in modeling the microstructure properties in deformed structures and relating these structures to the stress needed to deform the material. So, we have to design a system such that, given the values of dislocation cell wall thickness, cell size, and dislocation density, the system will be able
to predict the stress needed to deform the sample under these conditions given different strains (0%, 5%, 15%, 15%, and 20%). As a result, we can develop a relation between these strains and the stress under each strain.

We can treat this as a learning problem in which we have three input variables that have to be modeled in such a way as to predict the stress needed. We developed our system based on an ANN in which we have those three microstructures (dislocation cell wall thickness, cell size, and dislocation density) as inputs in the input layer. These inputs must be manipulated through hidden layer(s) to produce one quantity out of the output layer which is the stress. Once again, we have different strain values and rooming temperatures or hardening cases under which this system must be analyzed.

So, as a whole, our system has the following design and model procedure:

- Design the algorithm needed to compute the three different microstructure properties which are:
  
  → Dislocation Cell wall thickness
  
  → Dislocation Cell size
  
  → Dislocation Density

  One point to keep in mind is that we have the relations between the point-point distance and the misorientation angle and, using these relations, we have to develop the algorithms to extract those three parameters.

- Apply the above microstructure properties to our ANN as inputs
• Train the ANN until good results (stress values) with reasonable Mean Square Error (MSE) are achieved.

• The training process is as follows:
  → We have determined the number of hidden layers that must be used and the number of neurons in each hidden layer. Usually this process will be done by trial and error until a good prediction with reasonable MSE is achieved.
  → The learning rate ($\alpha$) is adjusted until good convergence occurs to the system and the prediction process succeeded. Usually this learning rate must begin at a high value (like 10) and then it must be decreased gradually until a convergence is obtained.
  → The weights that transfer the inputs to the hidden and output layers must be observed and saved, so we will be able to determine the average weight of each input and thus the input that has the highest impact on the result can be determined easily. This approach will be very helpful in analyzing the whole sensitivity of our system.

• Using the remaining set (say 15 out of 45 training examples) to test the system; this technique is referred to as cross-validation technique in which we will test our system against overfitting which is a common problem in learning problems because the system may seem to work very well under specific set of training examples, but when the system is tested using another example, it will fail to predict the correct value of the output. Strictly speaking, the system memorizes
particular training set and reacts well using this set but fails using a new set. More
details on cross validation and overfitting will be covered in Section 4.2.

- Drawing the necessary graphs which are used to compare the results under
different conditions and cases (i.e., hardening conditions and strains as well as the
impact of some inputs on the obtained result).
- Extracting a sensitivity table in which the average weight of each input of our
system must be computed to see which input has the highest impact on the output.
- Trying other learning methods like Support Vector Machine Learning (SVM) or
genetic algorithms (if possible).

Since the values of microstructure properties are not computed in advance, we need to
develop the required algorithms to compute their values. Once again, this will be done
using the relation between the point-point distance and the misorientation angle.

Figure 3.5 demonstrates a graphical representation of the input parameters using
EBSD.

3.2.1 **Dislocation Cell Wall Thickness algorithm**

Dislocation cell wall thickness means the width (thickness in microns) of the
boundaries that separates the grains into particular dislocation areas. For simplicity, it can
be referred to as the thickness of the lines between each pair of subcells that the material
consists of.
According to [22], dislocation in material science is a crystallographic defect, or irregularity, within a crystal structure. Our concern here is to develop an algorithm that computes the width of the boundary between each pair of those irregular structures as shown in Figure 3.5 given the relation between the point-point distance and the misorientation angle as mentioned above.

![Dislocation Microstructures](image)

\[ \sigma = f(D) \]
\[ D = g(\delta, d, p) \]

Figure 3.5 The goal is to develop a learning algorithm linking the dislocation structure parameters to the material properties.

The dislocation cell wall thickness algorithm works as follows:

1) Inputs the data set corresponding to a particular misorientation angle, strain level, and hardening condition.

2) Extracts the columns of point-point distances and the proper misorientation angles.
3) Goes through all the data points and stores the data points that have a misorientation angle $>15^\circ$ or $<1^\circ$ in a separate file. (This file must contain the complete details about the data points that don’t satisfy this condition, i.e., the point-point distance and the corresponding misorientation angle.)

4) Stores the location of the data points that don’t meet the above criterion in a separate array.

(i.e., those data points that have a misorientation angle $>15^\circ$ or $<1^\circ$).

5) Names this array as “index”. Thus, index (1) means the location (line) of the first point that doesn’t satisfy the above condition.

Now, to compute the values of dislocation cell wall thickness, we have four different cases:

*case 1*: If index (1) = 2 or if the location of the first point that doesn’t satisfy the condition is the second line in the data set then, the first value of thickness will be point-point distance (1) or the first value of the distance points. For simplicity, let’s assume that point-point distance is an array and it is referred to as D and the first element in that array is D (1).

*case 2*: Let’s assume that i refers to the location of the data point under study. If index (i+1) = index (i) +1, then skip this point and go immediately to 3 because in this case we have reached two consecutive points that don’t satisfy the misorientation angle condition.

*case 3*: If index (i+1) = index (i) +2 (i.e., there is one point between index (i+1) and
(index (i) +2) that does satisfy the condition, we will apply a simple trigonometric principle based on the triangles geometry as follows:

let \( t_0 = \text{index (i)} + 1; \)

\( t_1 = \text{index (i)} + 1 - 1; \)

\( t_2 = D(t_1) - D(t_1 - 1); \)

\( t_3 = M_{pp}(t_1) - 1; \) where \( M_{pp} \) is the array that contains the misorientation angles

\( t_4 = M_{pp}(t_1) - M_{pp}(t_1 - 1); \)

\( x_1 = D(t_1) - \frac{(t_2 \times t_3)}{t_4}; \)

\( t_5 = D(t_1 + 1) - D(t_1); \)

\( t_6 = M_{pp}(t_1) - M_{pp}(t_1 + 1); \) and

\( x_2 = D(t_1) + \frac{(t_5 \times t_3)}{t_6}; \)

Now, the value of thickness at \( i+1 \) location will be \( x_2 - x_1 \) (Figure 3.6).

Figure 3.6 The trigonometric operation to compute \( x_1 \) and \( x_2 \) in case 3
**case 4:** Otherwise, let’s assume that \( y = \text{index} \ (i+1) \) and \( z = \text{index} \ (i) + 1 \). Then the value of thickness at \( i+1 \) location will be \( D \ (y-1) - D \ (z) \) which is the difference of the point-point distance between the location \( z \) and the location \( y-1 \).

6) Repeat from step 3 until all the data points are analyzed.

### 3.2.2 Dislocation Cell Size Algorithm

Dislocation cell size means the width or the distance (in microns) between two points that are located on the boundary of a particular dislocation (subcell) area, as shown in Figure 3.5.

Thus, our concern is to develop an algorithm that computes the distance between any two points located in the boundary of each of those irregular structures given the relation between the point-point distance and the misorientation angle.

The dislocation cell size algorithm works as follows:

1) Inputs the data set corresponding to a particular misorientation angle, strain level, and hardening condition.

2) Extracts the columns of point-point distance and the proper misorientation angles.

3) Goes through all the data points and stores the data points that have misorientation angle >15° or <1° in a separate file. (This file must contain the complete details about the data points that don’t satisfy this condition, i.e., the point-point distance and the corresponding misorientation angle.)
4) Now the algorithm will store the location of the data points that don’t meet the above criterion in a separate array (i.e., the data points that have misorientation angle $>15^\circ$ or $<1^\circ$).

5) Names this array as “index” for example. Thus, index (1) means the location (line) of the first point that doesn’t satisfy the above condition.

Now, to compute the values of dislocation cell size, we have four different cases:

*case 1*: If index (1) = 2 or the location of the first point that doesn’t satisfy the condition is the second line in the data set then,

Let’s define a new array called avg which contains the average value of the point-point distance values that are located between the data points that don’t meet the misorientation angle condition.

So, in this case the first value in the avg array will be the same as the first value of thickness.

*case 2*: If index (i+1) = index (i) +1, then skip this point and go immediately to 3 because in this case we have reached two consecutive points that don’t satisfy the misorientation angle condition.

*case 3*: If index (i+1) = index (i) +2, let $t_0 = index (i) + 1$ and avg value at $i + 1$ location will be $D (t_0)$.

*case 4*: Otherwise, let’s assume that $y = index (i+1)$ and $z = index (i) +1$, then the avg value at $i + 1$ location will be the average of the points between the points $z$ and $y-1$. 
6) Repeat from step 3 until all the data points are analyzed.

7) Now, after computing the avg array, we have to go through this array using a loop as follows:

   for i =1 to (a2-1) where a2 is the number of elements in avg array
   
   cell\_size (i) = avg (i+1) - avg (i);

   where the cell\_size is the array that contains the dislocation cell size values. So, we notice that the difference of the two consecutive elements in avg value is the corresponding dislocation cell size value. Furthermore, note that if in any iteration of the loop the value of cell\_size is negative, we have to take the absolute value of it in order to make it positive.

   Figure 3.7 shows a graphical representation of the dislocation cell wall thickness and the cell size on point-point distance vs. misorientation angle plot. (Note: the graph is aggregated for simplicity).

Figure 3.7 A graphical representation of Dislocation Cell Wall Thickness and Cell Size
3.2.3 Dislocation Cell Wall Thickness and Cell Size Analysis

After describing the algorithms for extracting the values of dislocation cell wall thickness and dislocation cell size using the relation of point-point distance and misorientation angle, let’s go further and analyze the occurrence frequency of dislocation cell wall thickness and dislocation cell size when the misorientation angle is 45°, -45°, and 90°.

Our analysis is based upon the occurrence frequency of dislocation cell wall thickness and cell size when the misorientation angle is 45°, -45°, and 90°. This analysis is done under the three cases of peak-aged (8hr170), under-aged (room temperature age) and over-aged (7000hr170).

Figure 3.8 shows the frequency of dislocation cell wall thickness under “under-aged” conditions for three data sets that have misorientation angle of 45°, -45°, and 90° respectively.

Figure 3.9 shows the frequency of dislocation cell wall thickness under “peak-aged” conditions for three data sets that have misorientation angle of 45°, -45°, and 90° respectively, and Figure 3.10 shows the same type of analysis but for “over-aged” conditions.

Figure 3.11 shows the frequency of dislocation cell size under “under-aged” conditions for three data sets that have misorientation angle of 45°, -45°, and 90° respectively, Figure 3.12 shows the same analysis for the “peak-aged” case, and Figure 3.13 shows the same analysis for the “over-aged” case.
Figure 3.8 Frequency count for Dislocation Cell Wall Thickness values for three different misorientation angles under “under-aged” case

Figure 3.9 Frequency count for Dislocation Cell Wall Thickness values for three different misorientation angles under “peak-aged” case
Figure 3.10 Frequency count for Dislocation Cell Wall Thickness values for three different misorientation angles under “over-aged” case

Figure 3.11 Frequency count for Dislocation Cell Size values for three different misorientation angles under “under-aged” case
Figure 3.12 Frequency count for Dislocation Cell Size Values for three different misorientation angles under “peak-aged” case

Figure 3.13 Frequency count for Dislocation Cell Size values for three different misorientation angles under “over-aged” case
From Figures 3.8, 3.9, and 3.10 we conclude that the distribution of dislocation cell wall thickness values versus the frequency count of these values will be almost the same for the three sets that have a misorientation angle of 45°, -45°, and 90° respectively and this result is applicable for all hardening conditions illustrated above.

Furthermore, from Figures 3.11, 3.12, and 3.13 we conclude that also the distribution of dislocation cell size values versus the frequency count of these values will be almost the same for the three sets that have a misorientation angle of 45°, -45°, and 90° respectively and this result is applicable for all hardening conditions illustrated above.

Another analysis we have done is of the relation between dislocation cell wall thickness and dislocation cell size values versus the frequency count of these values for all of the three hardening cases under each strain.

Figures 3.14, 3.15, 3.16, 3.17, and 3.18 show this relation (using dislocation cell wall thickness) when the strain is 0%, 5%, 10%, 15%, and 20% respectively and Figures 3.19, 3.20, 3.21, 3.22, 3.23 illustrate this relation (using dislocation cell size) when the strain is 0%, 5%, 10%, 15%, 20% respectively.
Figure 3.14 The Frequency of Dislocation Cell Wall Thickness under three hardening conditions for 0% strain

Figure 3.15 The Frequency of Dislocation Cell Wall Thickness under three hardening conditions for 5% strain
Figure 3.16 The Frequency of Dislocation Cell Wall Thickness under three hardening conditions for 10% strain

Figure 3.17 The Frequency of Dislocation Cell Wall Thickness under three hardening conditions for 15% strain
Figure 3.18 The Frequency of Dislocation Cell Wall Thickness under three hardening conditions for 20% strain

Figure 3.19 The Frequency of Dislocation Cell Size under three hardening conditions for 0% strain
Figure 3.20 The Frequency of Dislocation Cell Size under three hardening conditions for 5% strain

Figure 3.21 The Frequency of Dislocation Cell Size under three hardening conditions for 10% strain
Figure 3.22 The Frequency of Dislocation Cell Size under three hardening conditions for 15% strain

Figure 3.23 The Frequency of Dislocation Cell Size under three hardening conditions for 20% strain
We have noticed from Figures 3.14-3.23 that there is a very slight difference in the relation of the frequency of dislocation cell wall thickness and dislocation cell size for each strain taking into account the three different hardening conditions that have been used with each of the five strains considered in the analysis (0%, 5%, 10%, 15%, and 20%).

3.2.4 Dislocation Density calculation technique

As referred to in [23], the geometrically necessary dislocation (GND) density is a microstructural physical quantity that is directly related to local non-uniform plastic deformations and is required to preserve the compatibility of the crystallographic lattice in cases of unevenly distributed plastic slip.

The GND density was determined to be the most important measured parameter affecting the yield stress. Experimental and statistical analysis showed a linear relationship between yield stress and average GND density [24]. So, GND density is determined to be the major microstructural parameter, sufficient to represent all characteristics of dislocation structures affecting the yield strength.

According to [24], it was observed that GND density by itself can sufficiently represent the strength contribution due to dislocation structures and samples with higher GND density observed higher flow stress.

In our model, the GND density values are provided in separate data sets, so we have a data set of point-point distance versus misorientation angle relation and another
data set for GND density and both of these data sets represent one training set for our model.

However, when the values of GND density are provided to our model using one of the sets, we have discovered that the number of values is too large (i.e., a multiple of 10000), so we need to find a way to reduce this number to a reasonable value in order for the system to afford the number of features in each input.

We have found out that the number of values of dislocation cell wall thickness and dislocation cell size is reasonable and acceptable, so we have developed a new technique that converts this huge number of GND density values to a number equal to the dislocation cell wall thickness and cell size values.

In this technique, we considered the GND values that we got initially from the data set as one cell. Then, we divided this cell into a number of subcells equal to the number of values in the dislocation cell wall thickness and dislocation cell size sets.

Each subcell contains a number of values depending upon the size of the original GND density cell and the number of subcells or the number of dislocation cell wall thickness and the cell size values as well.

Now, the average of each subcell’s values has been computed and thus the new averages have become the new set of GND density.

Figure 3.24 illustrates the whole “Cell” technique.
As we observe from Figure 3.24, the values of the original GND density matrix or dislocation density matrix are well-distributed among the subcells. This will give us a clear indication that the new values of GND density have a more aggregated and simpler format than the original GND density values.

In this way the number of values in each input of our model will be exactly the same and the system will be able to accommodate it as well.

3.3 Statistical Measures of Inputs Distributions

As mentioned in Section 3.2.4, the number of features (values) in each input of our ANN system (dislocation cell wall thickness, dislocation cell size, and GND density)
must be exactly the same, so the system can accommodate all of these values and produce the desired result of flow stress response.

However, since the number of features in each input is very large per dataset, there is a need to express the distributions of dislocation cell wall thickness, dislocation cell size, and GND density by the appropriate statistical measures by which we can summarize these features. Instead of having three inputs with too many features per input, we will have a significantly smaller number of inputs and the ANN system will work much better in terms of accuracy, architecture, and running time.

We have figured out that mean, standard deviation, and skewness are the most appropriate measures that our system could handle easily and they are applicable to all of our training and test sets under all of the hardening conditions we have considered (peak-aged (8hr170), under-aged (room temperature age) and over-aged (7000hr170)).

The mean is the average of values located within the distribution under concern and it can be expressed by the relation

$$\mu = \frac{1}{n} \sum_{i=1}^{n} X_i$$  \hspace{1cm} (3.1)

where $X_i$ is the values of the distribution under concern.

According to [28], the standard deviation of a probability distribution, random variable, or population or multiset of values is a measure of the spread of its values under a particular distribution. It can be expressed by:
\[ S = \left( \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \mu)^2 \right)^{1/2} \]  \hspace{1cm} (3.2)

Skewness is a measure of the asymmetry of the data around the sample mean. If skewness is negative, the data are spread out more to the left of the mean than to the right. If skewness is positive, the data are spread out more to the right. The skewness of the normal distribution (or any perfectly symmetric distribution) is zero [29].

The skewness of a distribution is defined as:

\[ Y = \frac{E(X - \mu)^3}{S^3} \]  \hspace{1cm} (3.3)

where \( E(t) \) represents the expected value of the quantity \( t \).

In this manner, we obtained three statistical measures, of mean, standard deviation, and skewness for each input of our ANN model. As a result, dislocation cell wall thickness features can be summarized by those three statistical measures and the same with the dislocation cell size and GND density values, resulting in 9 inputs to be applied to the ANN model representing the three quantities of the dislocation structures.

This type of analysis will avoid the burden of having too many features per input as well as enhancing the running time performance, generalizing the learning problem so that the model will work well with other datasets, and avoiding the overfitting problem (see Section 4.2 for details).
3.4 Model Architecture

As mentioned above, we used an ANN to model our system and we have found that the ANN gave us very reasonable results with almost negligible mean square error (MSE).

Once again, our model is able to predict the amount of stress needed to deform the material sample given the mean, standard deviation, and skewness of dislocation cell wall thickness, dislocation cell size, and dislocation density or “GND density” as inputs. Furthermore, the system has measured the required stress under four different percentages of strain, 0%, 5%, 10%, 15%, and 20%, for three different hardening conditions, peak-aged (8hr170), under-aged (room temperature age) and over-aged (7000hr170).

Each predicted stress value under particular strain represents the stress value needed to deform the sample that has 45°, -45°, or 90° of misorientation angle under this strain. So three different cases of misorientation angle values have the same amount of stress for the same percentage of strain.

After computing the mean, standard deviation, and skewness of all dislocation structures we have considered, we found out that those statistical measures are almost the same for material samples that have 45°, -45°, or 90° of misorientation angle under the same strain. So, it makes more sense to take the average value of each of those statistical measures, taking into consideration that this average value must be computed for the
samples that have misorientation angles of 45°, -45°, or 90° under the same strain as shown in Table 3.1.

In summary, after computing the average value of the statistical measures of different misorientation angles under the same strain in the cases of under-aged, over-aged, and peak-aged, we got 15 datasets as a whole and these datasets must be used for training and testing (validating) the ANN.

In general, the following factors must be taken into account in the design of any ANN system:

- Number of hidden layers.
- Number of neurons (nodes) operated in each layer.
- Mean Square Error (MSE).
- The weights of interconnection between neurons.
- Learning rate (α).
- Activation functions used (for input-hidden layers and for hidden-output layers).

Moore et al. [30] illustrate a technique that has been used to overcome a situation in which an ANN has only a small number of training examples, which is the case in our model. This technique is called “leave-one-out cross validation”. Thus, if we have 15 examples, in the first iteration 14 examples must be used for training the ANN model and the remaining example can be used for testing the model. Then, in the second iteration,

\[\text{\footnotesize \footnote{More details on cross-validation are mentioned in Section 4.2 Cross Validation Analysis}}\]
the other 14 examples must be used for training and the ANN should be tested using an example different than the example that was used in the first iteration to test the model. Again, in the third iteration, other 14 examples are used for training and the model should be tested using an example different than the examples that were used in the first and second iterations to test the model. This procedure will continue until we train the ANN model 15 times and test it using all the 15 examples that are available to make sure that it works well for external examples beyond the examples that are used for training.

On the other hand, even though 15 training examples are considered to be small in a typical learning problem, they seem to be enough in our model. It has been known that the more noise in the training data, the more examples we need in order to train the model and achieve the required generality and optimization. However, the calculation and extraction of the three statistical measures (mean, standard deviation, and skewness) based on the distributions of dislocation cell wall thickness, dislocation cell size, and dislocation density will significantly decrease the noise that may be in the original values of those three dislocation quantities. (On average, each group of 500 values that represents one of the three dislocation structure is converted to just 3 values of mean, standard deviation, and skewness, resulting overall in 9 applied inputs to the model.)

So, after training and testing our system on 15 training examples (15 training examples of point-point distance versus misorientation angle and 15 training examples that contain the values of GND density), we have obtained the following parameters for the ANN:
Table 3.1 Statistical analysis sample under “under-aged” condition; instead of having 15 datasets representing three
different orientations per strain, only 5 sets seem to be reasonable after computing the average value for all
statistical measures under the same strain rate. This is applicable to “over-aged” and “peak-aged” cases as well.

<table>
<thead>
<tr>
<th>Strain Rate (%)</th>
<th>Mis. Angle(°)</th>
<th>Dislocation Cell Wall Thickness</th>
<th>Dislocation Cell Size</th>
<th>GND Density</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Skewness</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>0.127</td>
<td>0.264</td>
<td>2.606</td>
<td>1.664</td>
</tr>
<tr>
<td>-45</td>
<td>0.136</td>
<td>0.259</td>
<td>2.139</td>
<td>1.905</td>
</tr>
<tr>
<td>90</td>
<td>0.087</td>
<td>0.224</td>
<td>2.937</td>
<td>1.417</td>
</tr>
<tr>
<td>Average Value</td>
<td>0.150</td>
<td>0.304</td>
<td>7.682</td>
<td>1.635</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>0.140</td>
<td>0.305</td>
<td>2.440</td>
<td>1.084</td>
</tr>
<tr>
<td>-45</td>
<td>0.158</td>
<td>0.331</td>
<td>2.942</td>
<td>1.229</td>
</tr>
<tr>
<td>90</td>
<td>0.189</td>
<td>0.380</td>
<td>2.412</td>
<td>1.199</td>
</tr>
<tr>
<td>Average Value</td>
<td>0.164</td>
<td>0.339</td>
<td>2.598</td>
<td>1.174</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>0.179</td>
<td>0.362</td>
<td>2.003</td>
<td>0.961</td>
</tr>
<tr>
<td>-45</td>
<td>0.211</td>
<td>0.404</td>
<td>2.333</td>
<td>0.932</td>
</tr>
<tr>
<td>90</td>
<td>0.180</td>
<td>0.376</td>
<td>2.523</td>
<td>0.888</td>
</tr>
<tr>
<td>Average Value</td>
<td>0.190</td>
<td>0.381</td>
<td>2.352</td>
<td>0.864</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>0.221</td>
<td>0.420</td>
<td>2.221</td>
<td>0.766</td>
</tr>
<tr>
<td>-45</td>
<td>0.248</td>
<td>0.450</td>
<td>1.969</td>
<td>0.920</td>
</tr>
<tr>
<td>90</td>
<td>0.238</td>
<td>0.463</td>
<td>3.068</td>
<td>0.641</td>
</tr>
<tr>
<td>Average Value</td>
<td>0.236</td>
<td>0.436</td>
<td>2.092</td>
<td>0.742</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>0.253</td>
<td>0.524</td>
<td>4.922</td>
<td>0.849</td>
</tr>
<tr>
<td>-45</td>
<td>0.278</td>
<td>0.504</td>
<td>3.064</td>
<td>0.871</td>
</tr>
<tr>
<td>90</td>
<td>0.258</td>
<td>0.460</td>
<td>1.977</td>
<td>0.758</td>
</tr>
<tr>
<td>Average Value</td>
<td>0.263</td>
<td>0.496</td>
<td>3.294</td>
<td>0.836</td>
</tr>
</tbody>
</table>
- 6 Neurons in one hidden layer.
- An output layer with one unit, so this means that the ANN model developed is doing regression analysis to predict just one single value of flow stress per training example. This is different from the classification analysis in which several categories of output are predicted.
- Learning rate \((\alpha) = 0.15\).
- Number of Epochs with most accurate output = 800.

Figure 3.25 illustrates the general architecture of the ANN we used in our model:

![Model Architecture Diagram](image)

Figure 3.25 Model architecture (9 inputs, one hidden layer of 6 neurons, and an output layer with one output. Learning Rate = 0.15). Each quantity of dislocation structure can be represented by 3 statistical measures of Mean, Standard Deviation, and Skewness.
CHAPTER IV

RESULTS OF TRAINING

In this chapter we will present the different analyses we have developed in our model as well as some graphs that show explicitly the model prediction of stress values versus the experimental values under four different strains; 0%, 5%, 10%, 15%, and 20% in three different conditions of peak-aged (8hr170), under-aged (room temperature age), and over-aged (7000hr170).

Once again, each predicted stress value under a particular strain will be the same for cases where the misorientation angle may differ. In our model the misorientation angle may take the values of 45°, -45°, and 90°.

4.1 Cross-Validation Analysis: Early Stopping method

The cross-validation or early stopping technique is used in a learning problem to make sure that the system will perform well and predict a correct value for all kind of input data sets in all conditions and cases.

In this technique, the training examples will be divided into two sets; a “training set” and a “validation set”. The validation set is used to determine when to stop training and the training should be stopped when the error on the validation set begins to rise due to overfitting or “overtraining”.
According to [26], overfitting is a common problem in machine learning that happens because the system seems to memorize specific training data and will not perform well on any other new data that may be given to the system in the future. Cross-validation techniques are used to avoid this phenomenon.

Sometimes, the training data is divided into a training set, a validation set, and a test set. The training set is used for training; validation set is used for detecting overfitting; and test set is for assessing performance.

Figure 4.1 [27] demonstrates the cross-validation technique which helps us in detecting overfitting by observing the relation of both the training set and the validation set versus the MSE over time or over the number of epochs. Once the MSE of the validation set begins to increase, we have to stop training since we have reached the point where overfitting occurs.

In our analysis, as mentioned in Section 3.4, we have 15 examples (data sets). Using the “leave-one-out cross validation” technique, we train the ANN model 15 times with 14 examples each time (the remaining set must be used for testing the model). As a result, 15 sets were used for testing (validation) to determine if overfitting occurs in the system.
Figure 4.1 Cross-Validation technique: the weights that result in the least MSE on the validation set are saved, so we can return to them after overtraining is detected.

Figure 4.2 shows a graphical representation of the performance of those 15 training times (with 15 sets each time) versus the performance obtained by 15 testing (validation) examples as mentioned earlier. This performance is measured using the relation between the number of epochs (running time) versus the MSE.

We notice from Figure 4.2 that the MSE in both the training and validation sets is almost negligible at 600 epochs and it will remain almost constant for both sets till 800 epochs of running time at which the lowest value of MSE is achieved. After 800 epochs the MSE begins to increase for the validation set and remains constant for the training set; this means that after 800 epochs overfitting was detected and the training process...
must be stopped at 800 epochs and all the weights and biases must be saved in order to use them for future sets because these weights and biases represent the best performance of our ANN model.

Figure 4.2 Cross-Validation analysis using a validation set and a training set

4.2 Model Prediction vs. Experimental values

As mentioned above, our model is used to predict the amount of flow stress (in MPa) needed to deform the material based on the values of dislocation cell wall thickness, dislocation cell size, and GND density, taking into account the statistical measures they represent (mean, standard deviation, and skewness).
This process must be done under different strains; 0%, 5%, 10%, 15%, and 20% for the hardening conditions of peak-aged (8hr170), under-aged (room temperature age), and over-aged (7000hr170) and as we know, each value of stress under a particular strain will be the same for the three different sets that have misorientation angles of 45°, -45°, and 90° respectively. This means that, regardless of the misorientation angle, the same value of flow stress will be predicted under a particular strain. That’s why the average values of mean, standard deviation, and skewness were computed for all three dislocation quantities under a particular strain no matter what the misorientation angle (see Sections 3.3 and 3.4).

Table 4.1 illustrates the predicted flow stress values for all 15 validation examples used in the model.

Figure 4.3 shows a complete view of our model with the predicted and experimental values taking into account the above conditions and cases.
Table 4.1 The predicted Flow Stress (in MPa) for each validation example used in ANN model

<table>
<thead>
<tr>
<th>Strain (%)</th>
<th>Dislocation Cell Wall Thickness</th>
<th>Dislocation Cell Size</th>
<th>GND Density</th>
<th>Flow Stress (MPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>std</td>
<td>skew</td>
<td>mean</td>
</tr>
<tr>
<td>0</td>
<td>0.117</td>
<td>0.268</td>
<td>2.561</td>
<td>1.635</td>
</tr>
<tr>
<td>5</td>
<td>0.164</td>
<td>0.339</td>
<td>2.598</td>
<td>1.174</td>
</tr>
<tr>
<td>10</td>
<td>0.19</td>
<td>0.381</td>
<td>2.352</td>
<td>0.864</td>
</tr>
<tr>
<td>15</td>
<td>0.236</td>
<td>0.438</td>
<td>2.092</td>
<td>0.742</td>
</tr>
<tr>
<td>20</td>
<td>0.263</td>
<td>0.496</td>
<td>3.294</td>
<td>0.836</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under-aged</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.125</td>
<td>0.284</td>
<td>2.597</td>
<td>1.204</td>
</tr>
<tr>
<td>5</td>
<td>0.154</td>
<td>0.329</td>
<td>2.359</td>
<td>0.954</td>
</tr>
<tr>
<td>10</td>
<td>0.166</td>
<td>0.355</td>
<td>2.555</td>
<td>0.778</td>
</tr>
<tr>
<td>15</td>
<td>0.221</td>
<td>0.481</td>
<td>3.388</td>
<td>0.871</td>
</tr>
<tr>
<td>20</td>
<td>0.219</td>
<td>0.433</td>
<td>2.696</td>
<td>0.821</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Over-aged</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.154</td>
<td>0.361</td>
<td>2.747</td>
<td>1.453</td>
</tr>
<tr>
<td>5</td>
<td>0.157</td>
<td>0.327</td>
<td>2.334</td>
<td>0.945</td>
</tr>
<tr>
<td>10</td>
<td>0.195</td>
<td>0.352</td>
<td>2.568</td>
<td>0.793</td>
</tr>
<tr>
<td>15</td>
<td>0.21</td>
<td>0.413</td>
<td>2.459</td>
<td>0.718</td>
</tr>
<tr>
<td>20</td>
<td>0.231</td>
<td>0.494</td>
<td>3.014</td>
<td>0.722</td>
</tr>
</tbody>
</table>
Figure 4.3 The Flow Stress versus deformation percentages or strains (0 to 20%) for the model prediction and experimental data (for peak-aged, under-aged, and over-aged samples).

Figure 4.3 shows that the predicted stress values are very close to the experimental data in each of the three cases of peak-aged, under-aged, and over-aged. This indicates that our ANN system is able to very accurately predict flow stress under different strains for three different hardening conditions.
Figures 4.4 and 4.5 show the MSE analysis versus the number of epochs and the number of neurons in the hidden layer respectively.

We have noticed from Figure 4.4 that the MSE will be very small (about 0.0003) when the number of epochs is $\geq 800$. This means that the system requires some time before reaching the optimal point. This is due to the fact that we have many features (values) per dislocation structure (an average of about 500 values) and so the model takes some time to compute the three statistical measures associated with each dislocation.
quantity. However, the overhead created by calculating the values of dislocation cell wall thickness and dislocation cell size is not the only burden the model should accommodate, but also computing the new values of GND density after applying the “cell” technique\(^2\) degrades somehow the whole performance.

Furthermore, Figure 4.3 and 4.4 show that the model is able to predict the flow stress value in the range of experimental measurements within an acceptable error of about 1.5%-2% (Table 4.2).

Table 4.2 Percentage errors for the three different cases that the model analyzed.

<table>
<thead>
<tr>
<th>Hardening Condition</th>
<th>Percentage Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak-aged</td>
<td>1.19%</td>
</tr>
<tr>
<td>Over-aged</td>
<td>1.24%</td>
</tr>
<tr>
<td>Under-aged</td>
<td>2.57%</td>
</tr>
</tbody>
</table>

This means that the parameters extracted from dislocation microstructures along with the statistical quantities of *mean*, *standard deviation*, and *skewness* are reasonable enough to forecast the flow stress response of the given material. The excellent agreement between model predictions and experimental results confirms that ANN modeling can provide a unique opportunity for materials modelers to study the materials behavior in complex structures and to develop more realistic structural based models. The information obtained herein supports existing dislocation hardening theories and offers a

\(^2\) For more details on “cell” technique, see Section 3.2.4
means by which mechanical properties can be determined from a detailed knowledge of the microstructure.

Figure 4.5 MSE versus the number of neurons in the hidden layer needed to achieve the optimum performance

We have noticed from Figure 4.5 that with 6 neurons in just one hidden layer, we will get the best performance with the least possible MSE. This relatively small number of neurons will accurately enable us to follow the operating weights that transfer the 9 inputs of our model to the hidden layer and from the hidden layer to the output layer.
Thus, the influence of each input in overall predicted flow stress can be easily analyzed and the generality of the ANN model on other external datasets will be guaranteed as well.

4.3 **Bayesian Regularization: Sensitivity Analysis**

In addition to the early stopping method introduced in Section 4.1, Bayesian regularization is another technique used to avoid overfitting in an ANN model especially when the number of training examples is small (only 15 training examples in our case). Matlab help documentation in [29] demonstrates that by using the appropriate training and performance functions, Bayesian Regularization technique determines the effective weights from the total weights that ANN originally has. Furthermore, Bayesian Regularization does not require that a validation set be separated out of the training set, but instead uses all the data. This means that the generality of the ANN model will be more likely to be achieved than the early stopping method.

The performance function used is MSEREG, defined as follows:

\[ MSEREG = \gamma \text{mse} + (1 - \gamma) \text{msw} \]  

(4.1)

where

\[ \text{mse} = \frac{1}{n} \sum_{i=1}^{n} (e_i)^2 = \frac{1}{n} \sum_{i=1}^{n} (t_i - a_i)^2, \]  

(4.2)

\[ \text{msw} = \frac{1}{n} \sum_{j=1}^{n} w_j^2, \]  

(4.3)
and $\gamma$ is the performance factor and in order to select its appropriate value, trial and error can be used or TRAINBR in Matlab can be used to determine the appropriate value automatically using the Bayesian Regularization routine in TRAINBR function.

Figure 4.6 shows our ANN architecture after applying Bayesian Regularization.

![ANN Architecture Diagram](image)

Figure 4.6 ANN model after applying Bayesian Regularization: out of 60 weights used, only 15 are effective; 9 from input to hidden layer (one per input) and 6 from hidden to output layer.
From Figure 4.6, we have seen that 6 neurons are needed in the hidden layer and the training is stopped when the sum squared error (SSE) and sum squared weights (SSW) are relatively constant over several iterations.

The values of flow stress obtained in all hardening conditions under all strains are the same as in using early stopping method introduced earlier. The MSE is about 0.0003 and the percentage error is about 2% and this is achieved after 1200 epochs.

However, the model implemented using Bayesian Regularization takes longer to converge than the early stopping method. But in our case, 1200 epochs of running time is still reasonable and it does not significantly degrade the whole performance compared to 800 epochs needed using the early stopping method.

We have also analyzed the influence of each of the nine inputs (three input groups of mean, standard deviation, and skewness corresponding to dislocation cell wall thickness, dislocation cell size, and GND density respectively) of our model on the predicted result of flow stress in each case of strain and hardening condition.

To know which input has the highest impact on the overall result, we can observe the weights that transfer the model from the input layer to the hidden layer and from the hidden layer to the output layer. The input with the largest weight has the highest influence on the system. Note that the absolute value of the negative weight must be applied to avoid the possibility of having a case where the positive values of weights cancel out the negative ones.

Table 4.3 shows the weight of each input of our model.
Table 4.3 The average weight of each input of our model. This indicates the level that each input affects the predicted result. Dislocation Cell Wall Thickness, Cell size, and GND density have average weights of 2.42, 3.12, and 3.24 respectively.

<table>
<thead>
<tr>
<th>Model Inputs</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dislocation Cell Wall Thickness (2.42)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.861</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.689</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.873</td>
</tr>
<tr>
<td>Dislocation Cell Size (3.12)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.194</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.764</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.161</td>
</tr>
<tr>
<td>Dislocation Density (GND) 3.24</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.917</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>1.527</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.800</td>
</tr>
</tbody>
</table>

From Table 4.3, we have seen that the impact of dislocation cell wall thickness, dislocation cell size, and dislocation density has the levels of 2.42, 3.12, and 3.24 respectively which means that the dislocation density input has the highest impact on our model with a weight of 3.24. This result agrees with [25] which illustrates that the GND density was determined to be the most important measured parameter affecting the yield stress, and experimental and statistical analysis showed a linear relationship between yield stress and average GND density. The parameter that has the second highest impact
on the output is the dislocation cell size and hence the dislocation cell wall thickness has the lowest impact on the overall predicted stress value.

Regarding the effect of individual inputs of mean, standard deviation, and skewness on overall predicted stress value, we have seen that all of them are useful in predicting the correct values of flow stress. The reason for this is that, as illustrated in Section 3.3, standard deviation is a measure of how the data points are spread under a particular distribution and, in our case, the main concern is to identify clearly the exact distribution of dislocation cell wall thickness, cell size, and GND density values along their frequencies all over the distribution area.

Skewness is also an important factor since it gives us an idea of where the data points are concentrated under the distribution. If those data points are concentrated in the left side of the distribution, this means that the relatively small values have much higher frequency than other values, while if the values are concentrated in the right side of the distribution, this means that the points that have larger values have much higher frequency than the rest of the values.

The mean gives us an idea about the average value of all data points in the distribution.

4.4 Multiple Linear Regression Analysis

Multiple Linear Regression is another method by which the learning problem can be modeled. But it assumes the relation between the inputs and the desired output(s) is
linear. In this method, there is no need to introduce an intermediate layer that transfers the inputs to it and from there to the output layer.

Typically, the multiple linear regression technique may work fine with simple systems where the relation between their parameters is linear. Such systems don’t require complex Artificial Intelligence techniques such as ANN, fuzzy logic, or genetic algorithms, as these methods may create unnecessary overhead that affects the performance and the running time of the model.

In addition, by using the linear regression method, we will avoid the risk of overfitting that may occur in other AI learning techniques, and the burden created by computing the operating weights (between the input and hidden layers and between the hidden and output layers in case of ANN) will be eliminated as well.

On the other hand, if the relation between the learning problem’s parameters is not linear, or there are cases that may arise in the future violating this principle, then this technique will not work as well as ANN.

In this Section, we will test the ability of multiple linear regression technique in modeling and predicting the desired values of flow stress response described in chapters III and IV.

The function used to apply the multiple linear regression technique to our model is called ‘regress’ and it is available as part of the MATLAB statistics toolbox as follows:

\[ b = \text{regress}(\text{Stress}, \text{Measures}) \]  

(4.4)
where ‘Stress’ is the 15 x 1 matrix that represents the flow stress values that the model (after applying multiple linear regression) is able to predict and ‘Measures’ is the matrix that includes the statistical measures needed to predict each flow stress value. The dimension of the ‘Measures’ matrix is 15 x 9 resulting in 9 dimensions (statistical measures) that represent the three dislocation quantities for all 15 training examples. Matrix ‘b’ is the matrix of all statistical measures’ coefficients.

Figure 4.7 shows the experimental values of flow stress (in MPa) under different strains (0%, 5%, 10%, 15%, and 20%) versus model prediction using multiple linear regression analyses in all three hardening conditions of peak-aged, over-aged, and under-aged. Our system cannot be modeled accurately using the linear regression technique because the relation between the inputs of dislocation cell wall thickness, dislocation cell size, and GND density (along with their statistical measures) and the flow stress is not linear, although this technique will give us an indication that over-aged points must be in the middle between the maximum and the minimum bounds that represent the peak-aged and the under-aged points respectively.
Figure 4.7 The prediction of Flow Stress values using multiple linear regression technique under the three considered hardening conditions in all strains. The points in the Figure (not filled) demonstrate the experimental values of Flow Stress obtained in each plastic strain percentage.

After applying the multiple linear regression technique, the relationship between the flow stress and the 9 inputs of our model can be expressed as:

\[
Stress = -2745.7X_1 + 1792.4X_2 - 59.5X_3 + 135.1X_4 - 193.6X_5 + 8.6X_6 + 0.3X_7 + 0.2X_8 - 5.9X_9
\]  

(4.5)
where:

$X_1$, $X_2$, and $X_3$ are the mean, the standard deviation, and the skewness of dislocation cell wall thickness quantity respectively.

$X_4$, $X_5$, and $X_6$ are the mean, the standard deviation, and the skewness of dislocation cell size quantity respectively.

$X_7$, $X_8$, and $X_9$ are the mean, the standard deviation, and the skewness of GND density quantity respectively.

The numbers illustrated in equation (4.5) are the coefficients of the 9 statistical measures mentioned above.

The average percentage error after applying this technique is about 15% with a MSE of about 0.0225, which is not accurate at all compared with the 2% error of ANN modeling with MSE of about 0.0003 as described in Chapters III and IV.

This analysis leads us to conclude that our model of dislocation material microstructures and flow stress response must be modeled using techniques that can model nonlinear relationships, such as ANNs.
A model based on feed-forward neural networks in simulating flow stress behavior of Al-Mg-Si alloys was proposed. The dislocation microstructure data extracted from the high resolution EBSD measurements were used to develop the necessary algorithms to compute dislocation cell wall thickness, dislocation cell size, and GND density values under particular strains (0%, 5%, 10%, 15%, or 20%) and misorientation angles (45°, -45°, or 90°). Then, those three dislocation quantities along with the statistical measures of mean, standard deviation, and skewness that represent each quantity were used to train an artificial neural network (ANN) model. The model captures the details of dislocation structure evolution by tracking the variation in Geometrically Necessary Dislocation (GND) density, dislocation cell size, and dislocation cell wall thickness.

The ANN model was robust enough to capture the structure-property relations in terms of predicting the flow stress of the investigated alloy that contained various precipitate morphologies within the range of experimental measurements with a maximum of 2.5%, error which is more accurate and effective than the multiple linear regression technique discussed in this study. This investigation offers a motivation to
utilize ANNs for modeling the complex dislocation structure evolution that is often too complex to model in a physical or mathematical based framework. The integration of this ANN model and finite element analysis will offer a unique opportunity for the materials science and solid mechanics community.

Furthermore, regarding the dislocation cell wall thickness and dislocation cell size analysis, we have found that they are almost the same for the misorientation angles of 45°, -45°, and 90° under particular strain, which means that these dislocation analyses are not misorientation dependent. This gives strong evidence that our model extracts reasonable results for dislocation cell wall thickness and dislocation cell size.

Another interesting observation is that regardless of the misorientation angle used under particular strain, the value of the predicted flow stress will be the same, taking into consideration that the analysis is done under the same strain. This means that it will make more sense if the average values of mean, standard deviation, and skewness are computed for all the dislocation quantities under a particular strain.

In this manner, fewer training examples were obtained and the biggest challenge was how to train and test the ANN model in order to achieve the required optimality and generality. A cross-validation technique known as “leave-one-out cross validation” and Bayesian Regularization were introduced as solutions that make the model converge to the desired result with a reasonable running time; MSE, and generalization to other datasets was guaranteed as well. However, by using the appropriate training and
performance functions, Bayesian Regularization minimizes the number of effective weights and hence simplifies the whole architecture of ANN.

The same kind of analysis was done for different hardening conditions (peak-aged (8hr170), under-aged (room temperature age) and over-aged (7000hr170)) under different strains, and it turned out that the number of dislocation cell wall thickness and dislocation cell size values are close to each other in different hardening conditions, although the analysis demonstrated that the number of dislocation cell wall thickness and dislocation cell size values under peak-aged condition is larger than that of over-aged, and the number of these dislocation values under over-aged conditions is larger than that of under-aged condition during each interval of values (i.e. between 0-1 microns, 1-2 microns, ……etc). These results are identical with the dislocation analyses that have been done in the literature.

After analyzing the weights that were obtained from our model, it was found that the dislocation density has the highest weight which means that it has the highest impact on the output of the model. In addition, among the statistical measures (of each dislocation structure) that were analyzed, it has been found that all of them are needed to predict the value of flow stress response accurately.
REFERENCES


[21] [http://www.mate.calpoly.edu/current/surface/techniques/ebsd.html](http://www.mate.calpoly.edu/current/surface/techniques/ebsd.html)


