Study on Ductility of Ti Aluminide Using Artificial Neural Network

R. K. Gupta, Rama Mehta, Vijaya Agarwala, Bhanu Pant, and P. P. Sinha

1 Materials and Mechanical Entity, Vikram Sarabhai Space Center, Trivandrum 695022, India
2 National Institute of Hydrology, Roorkee 247667, India
3 Department of Metallurgical and Materials Engineering, Indian Institute of Technology, Roorkee 247667, India

Correspondence should be addressed to R. K. Gupta, rohitkumar.gupta@vssc.gov.in

Received 25 March 2011; Revised 7 August 2011; Accepted 16 August 2011

1. Introduction

Ti aluminide has been an important aerospace material due to its high temperature properties and lower density as compared to superalloys. The ordered structure of aluminides that make them useful for high temperature applications makes them brittle at ambient temperature [1–3]. Therefore, inspite of having good properties, the usefulness of these alloys has been limited to some specific applications only. Room temperature tensile ductility is maximum (~1.5%) at around Ti-48Al (at%) aluminum, which is insufficient for further processing and applications. Hence, development of Ti aluminides has centered around Ti-48Al (at%) composition. It belongs to the γ (TiAl) plus α2 (Ti3Al) region of the phase diagram [4–6]. Various methods like alloy addition, controlled processing, heat treatment and so forth, are applied to get optimum combination of strength and ductility. Alloying additions in the range of 1 to 10 at% is studied with Cr, V, Mn, Nb, Ta, W, and Mo [6]. The alloying additions of V, Mn, Ni, and Cr in the range of 2–4 at% have shown enhancement in ductility of the alloy.

1.1. Phase Diagram

The Ti-Al phase diagram is shown in Figure 1. The γ (TiAl) plus α2 (Ti3Al) phase region is of primary interest for the study of Ti Aluminide. The phase diagram shows the stability of the γ (TiAl) phase up to 1370°C, above which the α2 (Ti3Al) phase becomes stable. The transformation of γ (TiAl) to α2 (Ti3Al) phase occurs by the formation of a η (Ti4Al3) compound at 1280°C.

1.2. Mechanical Properties

The mechanical properties of Ti Aluminide are shown in Table 1. The Young’s modulus, yield strength, and ultimate tensile strength of Ti-48Al (at%) are 180 GPa, 870 MPa, and 1070 MPa, respectively. The ductility of Ti-48Al (at%) is limited to ~1.5%, which is insufficient for further processing and applications.

Effect of microstructure on the mechanical properties was studied and duplex structure with fine grain size has been reported to be optimum for superior strength and ductility [4, 7–9]. To obtain desired microstructures and mechanical properties, the effect of several heat treatment cycles on aluminides has been studied at different temperatures and with varying cooling rates [9–16] and marginal improvement in ductility was reported. In this way several studies have been conducted with limited success in improving ductility of the alloy. However, experimental studies are expensive due to the use of high purity alloying elements and processing under controlled atmosphere. Here, theoretical models are very useful for optimization of process parameters. Experimentation with such optimized parameters shall minimize the number of experimental attempts and could lead to achieve desired ductility.

During the last decade, there has been an increased interest in applying new emerging theoretical techniques such as fuzzy inference system (FIS) and artificial neural network (ANN) for optimization-related problems [17–20]. These are the most common data driven models. These models...
intend to describe the nonlinear relationship between input (antecedent) and output (consequence) to the real system. In the present paper ductility of Ti aluminide is studied through ANN modeling.

Though ANN modeling is a relatively new technique, there has been an increasing interest in applying this technique in recent times in different fields of material science [21–25]. The basic advantage of employing ANN is that it does not require any external manifestation of parametric relationship. It learns from examples and recognizes patterns in a series of input and output values without any prior assumptions about their nature and interrelations. It consists of a number of interconnected computational elements called neurons. The neurons are arranged in three types of layers—input, hidden, and output. Information processing in a neural network occurs through interaction between these neurons. There is a wide range of ANN architectures [26], among which the three layer (input, hidden, and output) feed-forward architecture is used in the present work. It consists of layers of neurons, with each layer being fully connected to the preceding layer by interconnection weights $(W)$. A neural network representation with two inputs $x$ and $y$ is shown in Figure 1. Each of these input variables is associated with 3 neurons as decided after trial and error. The computations were performed in 5 levels (L1 through L5, which consists of 1 output, 1 input, and 3 hidden layers). Network with three hidden layers (Figure 1) was attempted but the error increased as well as computation time was more, therefore, it was restricted for single hidden layer network. The directions marked in Figure 1 imply the flow of information. The contribution of each rule toward the model output is computed in the fourth level (L4). Overall output of the model is computed at fifth level (L5) by combining the signals received from the previous level. Subsequently models were validated with the literature database using three important indices, that is root mean square error (RMSE), regression coefficient ($R^2$) and the model efficiency criterion of Nash and Sutcliffe [27].

![Figure 1: Architecture of artificial neural network system.](image)

### Table 1: Assumptions for categorizing literature data.

<table>
<thead>
<tr>
<th>Sl. no.</th>
<th>Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>Effect of alloying elements like Cr, V and Mn is same on ductility.</td>
</tr>
<tr>
<td>(2)</td>
<td>Effect of alloying elements on various alloys Ti 44-52Al is same as on Ti48Al</td>
</tr>
<tr>
<td>(3)</td>
<td>Grain size means diameter of grain for equiaxed grains and interlamellar spacing for lamellar grain.</td>
</tr>
<tr>
<td>(4)</td>
<td>Value of grain size for different composition follow the same trend as it follow for a specific composition and the trend is extrapolated.</td>
</tr>
<tr>
<td>(5)</td>
<td>Ductility data referred from literature is for mean grain diameter.</td>
</tr>
<tr>
<td>(6)</td>
<td>Ductility data for alloy type and grain size is the maximum ductility for the alloy in desirable heat treatment condition.</td>
</tr>
</tbody>
</table>

### 2. Ductility Parameters and Data for Model

Ductility values were consolidated from the literature [5–11, 28–33], and three important parameters, namely alloy chemistry, grain size, and heat treatment cycles were identified, which have major influence in ductility of the alloy. Alloy chemistry, grain size, and heat treatment cycles have important influence on ductility. For example, as content of aluminium increases in the alloy, ductility decreases. Binary alloy has limited ductility, and addition of ternary and quaternary alloying elements improves the ductility of the alloy. Similarly, finer grain size improves the ductility of the alloy through microstructural refinement avoiding segregation of impurities. Heat treatment cycle results in formation of desired phases, which has higher ductility.

Collected data were categorized in 2 types of alloy (2 components and multicomponent alloy), 4 types of grain sizes, and 7 types of heat treatment cycles with certain assumptions (Table 1). Ductility values collected from the literature with various categories of parameters are presented in Table 2, which were further interpolated according to trends in literature values and following the linearity basis. In this way 56 ductility data combinations were generated, in the same order as given in Table 2, for each alloy type and for each grain size in different heat treatment conditions. These values are termed as observed values. Total number of observed values is 56, in which dataset 1 is selected just by picking first 40 values for training and last 16 values for testing. In data set 2, first 10 values from 21 to 30 and from 41 to 50 values have been taken for training and the rest for testing. Data set 3 has picked randomly 43 values from each category for training data and similarly 30 values for testing data. Few data have been considered in both training and testing sets. Data selection for different data set has been given in Table 3. In the present paper, data extracted from the literature are written as observed data.
Table 2: Parameters, its detail characteristics, and ductility values.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Detail characteristics</th>
<th>Ductility (%ε)</th>
<th>Digital data form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alloy type</td>
<td>(1) Binary alloy (Ti-48Al at%) [5, 6, 28]</td>
<td>2.1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(2) Multicomponent alloy (Ti-48Al + 0–4 M*) [9, 11]</td>
<td>3.0</td>
<td>2</td>
</tr>
<tr>
<td>Grain size</td>
<td>(1) 10 μm [29, 31]</td>
<td>2.0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(2) 50 μm [29, 31]</td>
<td>1.8</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(3) 100 μm [32]</td>
<td>1.4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>(4) 250 μm [32]</td>
<td>1.1</td>
<td>4</td>
</tr>
<tr>
<td>Heat treatment cycles</td>
<td>(1) Above (T_α) (2 hrs) then FC to RT [32]</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(2) Just below (T_α) (20 min.) then AC to RT [11]</td>
<td>1.5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(3) Below (T_α) (100°C) soaking (2 hrs) then AC to RT [11]</td>
<td>3.6</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>(4) Above (T_α) (1 hr) then CC1 to below Te (soaking for 6 hrs) then AC to RT [8]</td>
<td>0.8</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>(5) Just above (T_α) (30 min) then CC2 to RT + heating below Te and soaking for 6 hrs then AC to RT [8]</td>
<td>2.8</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>(6) Just below (T_α) (3 hrs.) then CC1 to below Te (soaking for 4 hrs) then AC to RT [8]</td>
<td>3.8</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(7) Soaking just below (T_α) for 4 hrs then FC to ((T_α)-50°C) and soaking for 4 hrs then FC to RT + heating and soaking below Te (24 hrs) then AC to RT [33]</td>
<td>5.0</td>
<td>7</td>
</tr>
</tbody>
</table>

\(M^*\) is alloying elements (Cr/ Mn/V at the desired content ranging from 0–4 at%, leading to maximum ductility). \(T_α\) (α transus temperature for specific alloy), RT (room temperature), FC (furnace cooling), AC (air cooling), CC1 (controlled cooling 80°C/min), CC2 (controlled cooling 100°C/min), and Te (eutectoid temperature of alloy).

Table 3: Length of data used in different sets.

<table>
<thead>
<tr>
<th>Data sets</th>
<th>Training data</th>
<th>Testing data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset 1</td>
<td>40</td>
<td>16</td>
</tr>
<tr>
<td>Dataset 2</td>
<td>30</td>
<td>26</td>
</tr>
<tr>
<td>Dataset 3</td>
<td>43</td>
<td>30</td>
</tr>
</tbody>
</table>

3. Model Development

Three variables, namely composition, grain size, and heat treatment cycles were considered as antecedent and ductility as consequence during the modeling.

3.1. Formulation of ANN Model. Depending on the data availability, total data has been subdivided into training and testing sets as mentioned in previous section. The number of hidden levels, the number of input and output nodes and the number of nodes in the hidden levels were decided by trial and error method. A schematic diagram of a typical \(j\)th node is displayed in Figure 2. The inputs to such a node come from system variables or outputs of the other nodes, depending on the levels that the node is located in. These inputs form an input vector \(X = (x_1, \ldots, x_i, \ldots, x_n)\). The sequence of weights leading to the node form a weight vector \(W_j = (w_{ij_1}, \ldots, w_{ij_j}, \ldots, w_{ijd})\), where \(w_{ij}\) represents the connection weight from the \(i\)th node in the preceding level to this node.

The output of \(j\) node, that is, \(y_j\), is obtained by computing the value of function \(f\) with respect to the inner product of vector \(X\) and \(W_j\) as

\[
y_j = f(X \ast W_{ij} - b_j),
\]

where \(b_j\) is the threshold value, also called the bias, associated with this node. The function \(f\) is called an activation function. It can be determined with the response of a node with the total number of received input signals.

For selecting number of neurons, models with 3, 4, and 5 number of neurons were tried and RMS error were computed. Although it does not have much difference in RMS error for each type but least error was seen (Figure 3) for 3 neurons in all the cases. Also, with increasing number of neurons becomes more complex therefore three neurons were used in all the analysis.

3.2. Training (Learning) of ANN Model. To generate an output vector as close to the target vector, training process is employed. A network trains by adjusting the weights that link its neurons so as to find optimal weight matrices and
bias vectors that minimize a predetermined error function. It is written as

\[ E = \sum_{i=1}^{p} \sum_{j=1}^{q} (y_j - t_j)^2, \]

where \( t \) = component of the desired output; \( Y \) = corresponding ANN output; \( P \) = number of training parameters; \( q \) = number of output nodes.

A training data set is used to train the network or to determine the interconnection weights such that the response of the ANN clearly matches the observed behavior of the process being modeled. During training, typically mean square error (MSE) is monitored to find the optimal termination point for training. After training, the network is tested with the testing data set to determine how accurately the network can simulate the input-output relationship. If the performance of the ANN on the test data is satisfactory, the network is considered trained, weights are frozen and are used in actual application.

Different network processes have been tried with different sets of training and transfer functions to get optimum solutions. The network, which gives the optimum results with minimum error, could be frozen for testing or validation. The final frozen network contains “trainbr” as transfer function and “tansig” and “purelin” as training functions, as this set of combination with 3 neurons gives the optimum results. Accordingly, “tansig” and “purelin” training functions and “trainbr” transfer function have been fixed with trial and error for entire modeling. Number of experiments was carried out with all three data sets. The required ANN models (best ANN architecture) have been selected based on the training and performance checking of the models. Final architecture for all datasets is shown in Table 4.

3.3. Performance Indices. All models developed with different sets of data were tested and computed ductility was compared with literature-based ductility by means of RMSE (root mean square error) and \( R^2 \) (regression coefficient) statistics. The models performances were also evaluated by using Nash-Sutcliffe criterion [27] of percent variance (VAREX) as given below:

\[
\text{VAREX} = \left[ 1 - \frac{\sum_{i=1}^{N} (O_t - P_t)^2}{\sum_{i=1}^{N} (O_t - \bar{O})^2} \right] \times 100, \tag{3}
\]

where \( N \) = number of observations; \( O_t \) = observed value at time \( t \) (m\(^3\)/s); \( \bar{O} \) = mean of the observed values (m\(^3\)/s); and \( P_t \) = predicted ductility at time \( t \) (m\(^3\)/s). The value of VAREX ranges from 0 (lowest performance) to 100 (highest performance).

4. Results and Discussion

Artificial intelligence serves as a smooth interface between the qualitative variables and the numerical domains of the inputs and outputs of the model. As seen from the literature database, ductility of Ti aluminide is a function of number of variables, which are qualitative in nature. However, it can be presented mathematically where certain variables have primary effect on ductility of intermetallics and the same has been considered in the present work. The rule-base nature of the model allow the use of qualitative metallurgical relations expressed as general effect of different variables, which makes the model transparent to interpretation and analysis. Among the emerging modeling techniques ANN model has been considered to be very useful. Models are developed with the data generated from literature database with certain realistic assumptions and using different sets of data for training and validation. This makes the model more reliable.

Results of ANN models are presented in Table 5 and Figures 3–6. Figure 3 compares the RMS error for each model. It clearly shows that error is least for dataset 3. Similarly, Figure 4 compares the observed and achieved (modeled) ductility. Predicted ductility values are very close to observed values. Among the comparison of ANN results, data set 3 results are found to be very close to observed values. It also indicates achievable ductility value up to

**Table 4: Final architecture for all datasets.**

<table>
<thead>
<tr>
<th>Data set</th>
<th>SSE</th>
<th>SSW</th>
<th>Number of neurons</th>
<th>Number of epochs</th>
<th>Training functions</th>
<th>Transfer function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset 1</td>
<td>27.2377</td>
<td>20.602</td>
<td>3</td>
<td>143</td>
<td>Tansig</td>
<td>Trainbr</td>
</tr>
<tr>
<td>Dataset 2</td>
<td>20.2437</td>
<td>28.4964</td>
<td>3</td>
<td>114</td>
<td>Tansig</td>
<td>Trainbr</td>
</tr>
<tr>
<td>Dataset 3</td>
<td>0.7687</td>
<td>186.566</td>
<td>3</td>
<td>152</td>
<td>Tansig</td>
<td>Trainbr</td>
</tr>
</tbody>
</table>

SSE: squared sum of training error, SSW: squared sum of weights.
5.7% in specific combination of parameters, that is, for multicomponent alloy with grain size of 50 μm following heat treatment cycle number-7 (Table 2), that is, heating just below α transus temperature (Tα), soaking for 4 hrs → furnace cooling to Tα-50°C and soaking for 4 hrs at this temperature → furnace cooling to room temperature → heating below eutectoid temperature and soaking for 24 hours → air cooling to room temperature.

It is observed that models developed with data set 1 and 2 give results with high errors as compared to the models developed with data set 3. The minimum model error, that is, the difference between observed and modeled values gives the optimized parameters and it is found to be with data set 3. In Figures 5(a), 5(b), and 5(c) maximum error with individual dataset is presented, which is 0.36 for data set 3; 1.8 for data set 1; 2.2 for data set 2. It is purely due to the data selection. It is also noted that model developed with data set 3 gives best performance (Figure 6). It shows that if the input data covers as many data as under different conditions then it can develop an efficient model.

The performance of the model has been evaluated by three criteria as RMSE, regression coefficient, and variance (VAREX) shown in Table 5. It is clearly seen that RMSE is least for data set 3 and regression coefficient as well as VAREX...
Figure 6: (a, b, and c) Linear relationship between observed and modeled (ANN) ductility for dataset 1, 2 and 3.

Table 5: Performance Indices for different datasets using ANN technique.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>RMSE</th>
<th>VAREX</th>
<th>Regression coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset 1</td>
<td>0.254283</td>
<td>58.45</td>
<td>0.5926</td>
</tr>
<tr>
<td>Dataset 2</td>
<td>0.2276</td>
<td>87.90</td>
<td>0.3985</td>
</tr>
<tr>
<td>Dataset 3</td>
<td>0.04965</td>
<td>96.8</td>
<td>0.9691</td>
</tr>
</tbody>
</table>

is highest at 96.91% and 96.8%, respectively. It indicates that all the three performance indicators are close to ideal indices.

Scattering of data values is more with dataset 1 (Figure 5(a)) and least with data set 3 (Figure 5(c)). It also shows predicted ductility is close to observed reported values and it follows a definite trend. Therefore, use of different combinations of data sets for training and validation gives better prediction of property in ANN models. It means that model would be helpful in deriving optimum combination of parameters to obtain highest ductility in Ti aluminide intermetallics.

Analysis of the data has also been attempted through adaptive neuro fuzzy inference system (ANFIS) [19] using Takagi Sugeno model with subtractive cluster approach. Through ANFIS models, prediction accuracy is found to be depending on number of input variables with extent of data mixing [19]. Therefore, it can be inferred that selection of input data decides the output accuracy in ANFIS as well as in ANN models. Prediction from ANFIS approach [19] and present ANN approach is found to be supporting each other. It is also true that, if scattering of data is more, ANN models are more suitable, since it generates output by adjusting the interconnections between the layers. In addition to that, ANN models have less dependency on modeling parameters as noted in ANFIS models [19], so prediction accuracy shall be certainly better with ANN models.

5. Conclusions

(1) Models developed through ANN are giving results close to the observed reported ductility value especially with a wide range of data.

(2) It gives very good result with data set 3 when data are randomly selected from entire data, which covers wide range of data.

(3) Error (difference between observed and modeled values) is maximum for data set 2 as 2.2, while it is least as 0.36 for data set 3.

(4) Regression coefficient is high with dataset 3 in ANN modeling (0.969).

(5) Predicted ductility values are very close to observed values. Here achievable ductility value is obtained as 5.7% in specific combination of parameters. ANN predicts one specific combination. It is multicomponent alloy with grain size of 50 μm following heat treatment cycle number-7, that is, heating just below α transus temperature (Tα)→ soaking for 4 hrs → furnace cooling to Tα-50°C and soaking for 4 hrs at this temperature → furnace cooling to room temperature → heating below eutectoid temperature and soaking for 24 hours → air cooling to room temperature.

Acknowledgment

The authors extend their thanks to the Director, VSSC, for permitting to publish this paper.
References
