Association of Metallurgical Engineers of Serbia *Scientific paper*
AMES *UDC:669.715'782.018.11.001.573=20*

PREDICTION OF THE ALUMINUM SILICON MODIFICATION LEVEL IN THE AlSiCu ALLOYS USING ARTIFICIAL NEURAL NETWORKS

R. FRANCIS* and J. SOKOLOWSKI**

** ASQ Certified Reliability Engineer (2007), Senior Quality Engineer, Manufacturing Technology Group 3M Canada Company, 1840 Oxford Street East, London,ON,N5V 3R6 , Malicode: F2B-118 Email:rfrancis2@mmm.com, **Department of Mechanical Automotive & Materials Engineering, University of Windsor, Canada*

ABSTRACT

In this paper, two feed forward neural network models have been presented to predict the Silicon Modification Level (SiML) of W319 aluminum alloys using the Thermal Analysis (T.A) parameters as inputs. The developed neural networks are a Multilayer Perceptron (MLP) network and a Radial Basis Function (RBF) network. The neural network models were found to predict the SiML accurately (R=0.99). The accuracy of the Neural Network Models has been compared with the existing ΔT method and a linear multiple regression model. The comparison of the RBF and MLP networks has shown that the RBF requires much lesser training time than MLP.

Key words: modification, Al-Si-Cu alloys, artificial neural network

INTRODUCTION

W319 Aluminum is a casting alloy from the hypoeutectic Al-Si-Cu family. This alloy is mainly used in casting engine blocks, cylinder heads, and manifolds. Silicon is the most important alloying element in this alloy. It improves the castability and machinability of the W319 Al alloy. The unmodified W319 Al alloy contains silicon phase in the form of large plates with sharp sides and ends (acicular silicon). Modification is a method by which inoculants in the form of master alloys are added to aluminum melt to change the structure of the eutectic silicon from acicular flakes to a fine fibrous structure. Modification improves ductility and fracture strength of the aluminum alloys. Several modifiers are known (eg. Sr, Sb, Na, Ba, and Ca) of which strontium (Sr) is the one most commonly used in the industry owing to the following reasons: (a) It is easy to handle (b) It has good modification rate (c) It has long incubation period and (d) Its fading effect is low (Djurdjevic et al., 2001, Closset and Gruezleski, 1990).

The on-line prediction of the Silicon Modification Level (SiML) is very important for the quality control of W319 Aluminum alloy castings. Conventionally, SiML is

estimated by destructive sampling and manual microscopic analysis of the structure. The structure is compared with the American Foundry Society (AFS) chart for microstructure control in hypoeutectic alloys. In this chart the entire range of the modification level is divided into six classes based on the size of the eutectic silicon particles. Fully modified structure falls into class 5-6, partially modified into 2-4 and unmodified into class 1. Based on the visual comparison, the specimen is assigned a modification level, where the structures of the specimen and AFS chart are closest. But this method is expensive, time consuming, subjective and hence inaccurate. Also the microscope is not suitable to be used in the foundry environment as an online method to predict the SiML.

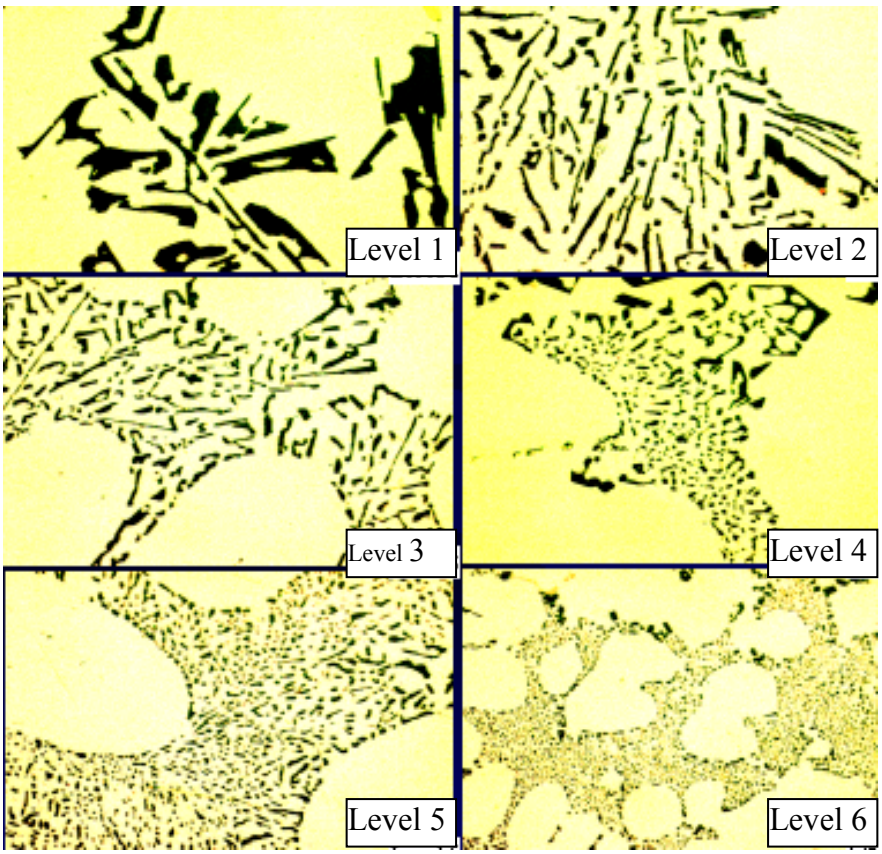

Fig.1. Representative Si structure morphologies corresponding to the AFS chart for hypoeutectic aluminum alloys (Djurdjevic et al., 2001).

Djurdjevic et al., 2001, presents another method to assess the SiML using the Thermal Analysis (T.A) technique. In this method the ΔT parameter of the cooling curve is correlated to the SiML estimated via Image Analysis (IA). ΔT parameter is the

difference in the Al-Si eutectic growth temperature between the unmodified and modified alloys. ΔT parameter can be obtained from the cooling curve using equation 1 (Djurdjevic et al., 2001).

$$\Delta T = T^{Al-Si}_{E,G,UNMODIFIED} - T^{Al-Si}_{E,G,MODIFIED} \quad (1)$$

where: $T^{Al-Si}_{E,G,UNMODIFIED}$ = aluminum-silicon eutectic growth unmodified temperature is the maximum temperature achieved during the recalescence of the melt from the nucleation temperature without any addition of modifiers or with a residual level of modifiers (8 ppm Sr).

$T^{Al-Si}_{E,G,MODIFIED}$ = aluminum-silicon eutectic growth modified temperature is the maximum temperature that is achieved during the recalescence of the solidifying melt.

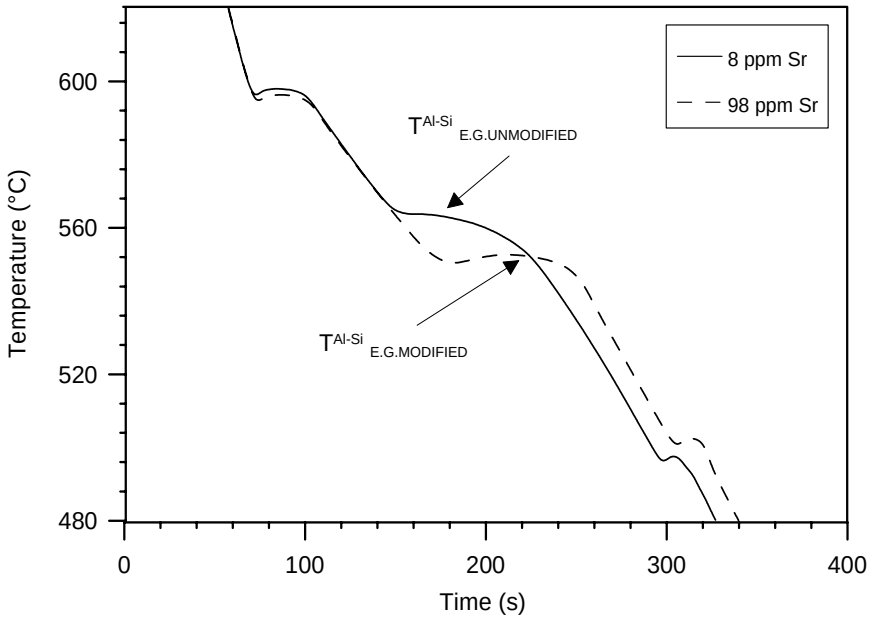

Fig.2. Cooling curves for low (8 ppm) and high (98 ppm) levels of strontium (Djurdjevic et al., 2001)

$$SiML = \frac{-17.0741}{\{1+\exp[(\Delta T - 13.42472)]/2.04788\}+18.06746} \quad (2)$$

From the knowledge of the ΔT parameter the SiML can be calculated using equation 2 (Djurdjevic et al., 2001).

This method improves the precision and accuracy of the SiML assessment because it eliminates the operator bias and subjectivity. This method gives a digital criterion for the assessment of SiML based on the AFS concept and can be used as an online method for the estimation of SiML.

In this paper, Neural Network modeling has been attempted to predict the SiML using the T.A. parameters as inputs. The on-line prediction of SiML is important for controlling casting microstructure and hence mechanical properties. It has been shown that TA can be used to predict the SiML (Djurdjevic et al., 2001). But the existing method also needs information about the unmodified alloy. So it would be worthy to develop a model to assess the SiML, only by just using TA parameters of the cooling curve. Artificial Neural Networks (ANN) has the potential to develop such a model owing to the following reasons:

1. ANN has adaptive learning algorithms and a trained neural network can generalize well on the test data. In other words ANN has self-learning and self-tuning capacity, it enables the output to approach successively to desirable output (Pospichal et al. 1997, Chen et al. 2002).
2. ANN can capture the interactions between the inputs because the hidden units are non-linear. The nature of the interactions are implicit in the values of the connection weighs (Bhadeshia, 1999). In this problem the interactions between the alloying elements are not well defined and the input-output relationship is complex and non-linear.
3. Modeling using ANN has been extensively used in the field of Materials Science and Engineering (Li et al., 1999, Song et al, 1995, Malinova et al., 2001, and Malinov et al., 2001).

ARTIFICIAL NEURAL NETWORKS (ANN)

BASIC PRINCIPLES

ANN consists of simple synchronous processing elements, which are inspired by biological nervous system (Calcaterra et al.2000). They solve a problem by means of learning rather than by specific programming based on well-defined rules (Lehr et al., 1994 and Beale et al. 2002). The input and output data are linked using a particular set of nonlinear basis functions. The basic unit of the ANN is the *neuron*. Neurons are connected to each other by links known as s*ynapses*; associated with each synapse there is a *weight factor*. Usually neural networks are *trained* so that a particular set of inputs produces, as nearly as possible, a specific set of target outputs (Malinov et al., 2001).

In feed forward ANN the information is processed in one direction - from input to output - and the neurons are arranged in layers. The numbers of input and output parameters determines the number of neurons in the input layer and the output layer respectively.

The neural network modeling involves the following steps:

a. Database collection.

b. Analysis and preprocessing of the data.

c. Training of the Neural Network. (This includes choice of the architecture, training functions, training algorithms and parameters of the network).

d. Testing of the trained network.

Experimental Procedure for Database Collection

The chemical composition of the commercial W319 aluminum alloy ingots used in this study is given in Table 1.

Table 1.Average chemical composition (wt%) of the W319-aluminum alloy used in the experiments

Si	Fe	Cu	Mn	Mg	Ni	Zn	Ti	Sb	P	Na
7.554	0.394	3.452	0.238	0.236	0.008	0.009	0.122	0.052	0.0015	0.0004

A total of 10 kg of W319 alloy ingots were melted in an electric furnace kept at a temperature of 735 ± 5 ^{0}C. Degassing was done for 15 minutes using an argon rotary degasser. The melt was modified through the addition of master alloy (Al-10 -wt% Sr). An incubation time of 15 minutes was allowed after strontium addition. The Strontium level was varied between 8 and 96 ppm (1 ppm=0.0001 wt%). Thermal Analysis (TA) was conducted using a system that records the temperature history of a 160 ± 10 g test sample of molten alloy as it cools down from the ~730^{0}C to 400^{0}C. The TA test samples were taken by submerging a cylindrical graphite cup (40 mm diameter, 50 mm deep) into the melt. The cup was kept submerged in the melt for 20s after which it was placed on the stand. Two K-type thermocouples, sheathed in stainless steel, were introduced into the cup to measure the temperature of the melt as it is cooled. The data for TA was collected using a high-speed data acquisition system linked to a personal computer.

Specimens for Image Analysis (IA) were cut from the test samples close to the tips of the thermocouples. The cross-sections of the specimens were ground and polished using standard metallographic procedures. The final polish was done using OP-U on microcloth. A Leica Q5501W Image Analysis System (IAS) was used to assess the Silicon Modification Level (SiML) of the polished test samples. Forty seven cooling curves (data sets) are obtained from the TA experiments with different levels of strontium ranging from 8 ppm(unmodified) to 96 ppm (fully modified).

Input Parameters of the ANN

The input parameters of the ANN were selected from the cooling curve parameters. The cooling curves were processed and the TA parameters (time, temperature, and fraction solid) were computed. The input parameters for the ANN were selected from the computed TA parameters by Stepwise Regression Analysis. The symbols of the selected TA parameters, their description, metallurgical meaning, and p-values are shown in the Table 2.

The p-value for a parameter tells whether the effect of that parameter is significant on the response (SiML). If p-value is less than or equal to the α-level (0.05) selected,

then the effect of the parameter is significant. If p-value is larger than the α-level, the effect is not significant (Devore, J.L., 1999). The selection of the input parameters is a very important aspect of ANN modeling. Usually the choice is based on the physical background of the process. All relevant input parameters must be included as inputs of the ANN. However those parameters that are linearly dependent should not be included as inputs.

Figure 3 shows the selected input parameters of the ANN in the cooling and fraction solid curves.

Table 2. Description of the Input parameters of the ANN

Symbol of the TA Parameter	Description	Metallurgical Meaning	p-value
T^{DEN}_{COH}	Temperature at the Dendrite Coherency Point (DCP)	At this point the dendrite tips of neighboring grains come in contact becoming fixed at their locations forming a skeleton throughout the sample. This point marks the transition from mass feeding to interdendritic feeding.	0
fs^{DEN}_{COH}	Fraction Solid at the Dendrite Coherency Point (DCP)	At this point the dendrite tips of neighboring grains come in contact becoming fixed at their locations forming a skeleton throughout the sample. This point marks the transition from mass feeding to interdendritic feeding.	0.012
$T^{Al\text{-}Si}_{E,NUC}$	Temperature at the Al-Si Eutectic Nucleation	As dendrites nucleate in the samples, the composition of the remaining liquid reaches the eutectic composition and the Al-Si eutectic begins to nucleate and grow.	0
$T^{Al\text{-}Si}_{E,G}$	Temperature at the Al-Si Eutectic Growth	Point at which considerable eutectic growth occurs. These parameters are used to gauge the morphology of the Al-Si structure.	0
$fs^{Al\text{-}Cu}_{E,NUC}$	Fraction solid at the Al-Cu Eutectic Nucleation	Start of the formation of the Al-Cu eutectic as the remaining liquid becomes enriched with Cu and Si. The temperature and duration of this phenomenon can be used to establish adequate solution treatment parameters.	0.016
$fs^{Al\text{-}Cu}_{E,MIN}$	Fraction solid at the Al-Cu Eutectic Nucleation	Point after stable co-precipitation of Al and Cu when latent heat generation equals the heat loss of the sample and appears as a minimum in the Al-Si region of the cooling curve.	0

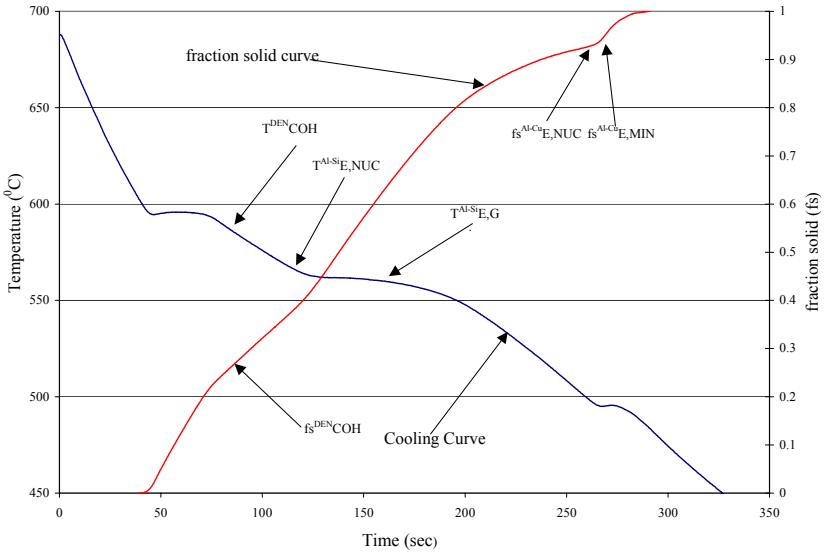

Fig.3. The input parameters of the ANN displayed in the Cooling and Fraction solid curves

Analysis and preprocessing of the data

The relationship between SiML and the input parameters has been established by performing a multiple linear regression. The regression equation and the R^2 value are given in equation 3.

$$SiML = 520 - 29.4fs^{DEN}{}_{COH} - 0.645T^{DEN}{}_{COH} + 0.165T^{Al\text{-}Si}{}_{E,NUC} - 0.238T^{Al\text{-}Si}{}_{E,G} - 49.9fs^{Al\text{-}Cu}{}_{E,NUC} - 45.3fs^{Al\text{-}Cu}{}_{E,MIN} \quad (3)$$

$$R^2=0.872$$

The high R^2 value indicates good fit between input parameters and SiML. The statistical significances (p-values) of these parameters are given Table 2.

The experimental data set is divided into training and testing sets. About 15% of the experimental data are used for testing the model. The test data set was so chosen that it is representing all the Modification Levels of the experimental data set (1 to 5). The training program is so written that each time when the program is run it starts with a different random distribution of the training data (Malinov et al., 2001). Prior to training, the data set is subjected to Principal Component Analysis (PCA). This procedure eliminates the highly correlated (redundant) input variables from the data (Beale and Demuth, 2001).

Neural Network Training

The software used to develop the model is the Neural Network Toolbox of MATLAB® 6.1. The process of fitting the network to the experimental data is called

training. It consists of adjusting the weight of each connection (*synapse)* between neurons. The weight of a synapse multiplied by the strength of the signal on that synapse gives the contribution of that synapse to the *activation* of the neuron for which it is input. The total activation of a neuron is then the sum of the activations of all of its inputs and this defines the value of the output signal for that neuron, via a *transfer function*. Transfer functions are generally s-shaped (sigmoid) curves, with the output value confined within the limits such as (0,1) or (-1/2,1/2). The most popular transfer functions are hard limit, linear, log sigmoid, and hyperbolic tangent sigmoid and radial basis function (Malinov et al., 2001).

By adjusting the values of synaptic weights throughout the network, the outputs of the ANN for any given set of inputs can be altered. During the training the weights are adjusted until the network correctly simulates the behavior of the system to be modeled. The simulation will rarely be exact; training is usually aimed at minimizing the sum of squares of the differences between predicted and measured values of the outputs. In this research two alternative neural networks were developed to predict the SiML. They are: (a) Multilayer Perceptron (MLP) Neural Network and (b) Radial Basis Function (RBF) Neural Network.

Multilayer Perceptron Neural Network

MLP is a feed forward network composed of an input layer, an output layer, and a number of hidden layers. The most popular algorithm for training of the MLP is the Backpropagation. In this algorithm a random initial set of weights are assigned to the network, then presenting the data inputs, comparing the network output to the desired output and adjusting the weights so as to reduce the corresponding error. This procedure is repeated until an acceptably low value of the error is achieved (Malinov et al., 2001). In this work the MLP is trained using the Levenburg-Marquardt backpropagation (Beale et al. 2002) algorithm. This algorithm is based on a standard numerical optimization technique. When a network is able to perform as well on test set as on training set inputs, the network is said to generalize well. In this work the generalizability of the MLP network has been attempted by means of regularization (Foresee and Hagan, 1997). Different network architectures have been investigated to determine the network that provides lesser time of training, and minimum generalization error. The architecture of the selected MLP network is given in Table 3.

Table3. Architecture of the MLP Neural Network

Number of hidden layers	Number of neurons in the hidden layer(s)	Number of neurons in the output layer	Transfer function of the hidden layers	Transfer function of the output layer	Training algorithm
2	6 in each layer	1	log sigmoid	linear	Bayesian regularization in combination with Levenberg-Marquardt training

Radial Basis Networks

This network was introduced to solve real multivariable interpolation problems. The Radial Basis Function (RBF) network consists of one input layer, one hidden layer and one output layer. The neurons in the hidden layer do not use weighted sun of inputs and sigmoid transfer functions, which are typical of MLP networks. Instead, the vector distance between the input weight vector and input vector determines the outputs of the hidden layer radial basis neurons. The output layer consists of linear neurons and produces a weighted sum of the outputs of the hidden layer (Beale and Demuth, 2001). The neurons in the RBF network have localized receptive fields, where as the sigmoid transfer function of the standard Multilayer Perceptron (MLP) networks creates global response (Campos-Velho et al., 1999). The RBF network trains faster than multilayer networks, but requires many neurons for high-dimensional input spaces (Beale et al. 2002).

The transfer function for a radial basis neuron is given by equation 4:

$$radbas(n) = e^{-n^2} \tag{4}$$

where n is the net input to the radial basis transfer function which is the vector distance between the input weight vector and input vector, multiplied by the bias vector. The training of RBF consists of deciding the number of hidden neurons, and the centers and spread of the transfer function and then training up the output layer (Campos-Velho et al., 1999). These decisions are made empirically by examining the vectors in the training data (http://web.odu.edu/engr/speechlab/paper2.pdf). The output layer weights are then trained using backpropagation (Campos-Velho et al., 1999). Radial basis networks tend to have many times neurons than a comparable MLP feed-forward network with sigmoid neurons in the hidden layer. Designing a radial basis network takes much less time than training a sigmoid/linear network (Beale and Demuth, 2001). The architecture of the developed RBF network is given in Table 4.

Table 4. Architecture of the RBF network

Number of hidden layers	Number of neurons in the hidden layer	Number of neurons in the output layer	Transfer function of the hidden layers	Transfer function of the output layer	Spread Constant
1	38	1	Radial Basis	Linear	2.3

It can be seen that the number of neurons in the hidden layer of RBF networks is much larger than that of the MLP (6).

Testing of the Trained Neural Network

The performance of the MLP network on the test data set has been evaluated. The net work was found to generalize well with the test data. Figure 4 shows the performance of the MLP on test data set.

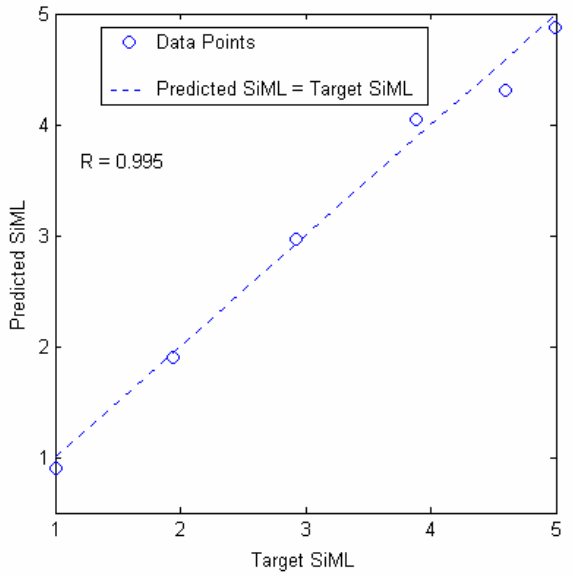

Fig.4. Performance of the MLP network on the test data.

Performance of the RBF neural network on the test data set is shown in Figure 5.

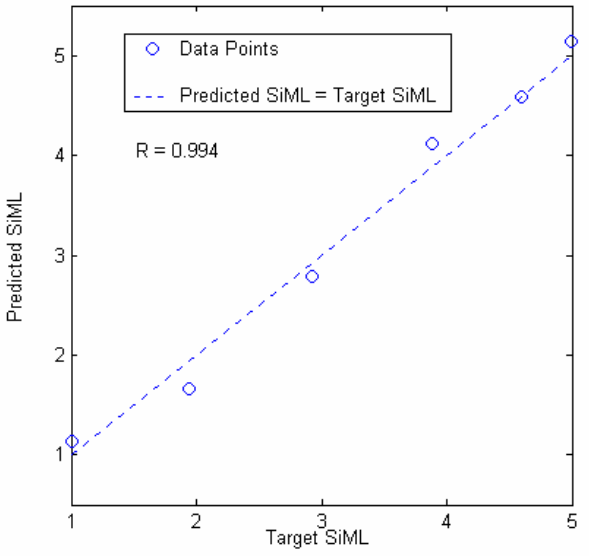

Fig.5. Performance of the RBF network on the test data.

The RBF network is also found to generalize well with the test data set.

RESULTS AND DISCUSSION

The Mean Square Error (MSE) of the test data set is the criterion used to evaluate the accuracy of the Neural Network. Lower MSE implies higher accuracy. MSE is calculated using the equation 5:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \qquad (5)$$

where:

N= size of the testing data set.

t_i = desired output value for the i^{th} input data.

a_i = predicted output value for the i^{th} input data by the model.

The MSEs of the MLP and RBF neural networks on the test data set has been compared with that obtained for the ΔT method and the Multiple Linear Regression. These results are shown in Table 5.

Table 5. Comparison of the Neural Networks with Regression, and ΔT method.

	MLP	RBF	Multiple Linear Regression	ΔT method
MSE	0.02	0.03	0.12	0.002

It can be seen that the mean square error is the least for the ΔT method. However, this method needs information about eutectic growth temperature of the unmodified alloy. The advantage of ANN approach is that it just requires the TA parameters and doesn't need any other information. But it still provides very high accuracy. The two feed forward neural networks developed in this project have been compared in terms of the MSE on the test data and the training time. Refer Table 6.

Table 6. Comparison of the Feed Forward Neural Networks.

Neural Network	MSE	Training Time (Seconds)
MLP	0.02	20.77
RBF	0.03	2.38

It can be seen that the MLP gives the lowest MSE. However, it should be noted that since MLP starts with random initial weights each time when the program is run, so that the MSE will not be the same lowest value every time. For, RBF network the MSE of the test data set will be the same every time when the program is run. The training time of the RBF network is much lower than that of the MLP network.

CONCLUSIONS

Two feed forward neural network models have been successfully developed to predict silicon modification level of the W319 Al alloys using the thermal analysis parameters as the inputs. The developed networks are a Multilayer Perceptron (MLP) network with two hidden layers and a Radial Basis Function (RBF) network with one hidden layer. The networks were found to generalize well with the test set data. The performances of the neural networks were compared with the ΔT method, and a Multiple Linear Regression Model. It has been found that the Neural Network models could accurately predict the SiML, just by using the TA parameters as the input. This shows that ANN in conjunction with TA can be used as a tool for the on-line prediction of SiML. The comparison of the MLP and RBF networks shows that the training time of the RBF network is much lower than that of the MLP network.

REFERENCES

[1] Djurdjevic, M., Jiang, H., Sokolowski, J.H., "On-line Prediction of Aluminum - Silicon Eutectic Modification Level Using Thermal Analysis", Materials Characterization, Vol.46, pp.31-38 (2001).

[2] Closset, B.M., Gruzleski, J.E., *The Treatment of Liquid Aluminum – Silicon Alloys*, American Foundrymen's Society, Illinois,IL (1990).

[3] Pospichal, J., Svozil, D., Vladimir, K., " Introduction to Multi-layer-feed-forward networks", Chemometrics and Intelligent Laboratory Systems, Vol.39, pp. 43-62 (1997).

[4] Chen, C., Wang, Z.P., Xu, J.L., "Neural Network Model in Predicting Mechanical Properties of Ductile Iron", AFS Transactions, 02-013, pp. 1-7 (2002).

[5] Bhadeshia, H.K.D.H., "Neural Networks in Materials Science", ISIJ International, Vol. 39 No.10, pp. 966-979 (1999).

[6] Li, D., Liu, Y., Meng, F., Zhang, Y., "Study on Thermal Analysis Models Used in Grey Cast Iron Quality Prediction", International. Journal of Cast Metals Research, 11, pp. 391 –394 (1999).

[7] Song, R., Tseng, M., Zhang, B., Zhang, Q., "The Application of Artificial Neural Networks to the Investigation of Ageing dynamics in 7175 Aluminum Alloys", Materials Science and Engineering: C, Vol.3, pp.39-41 (1995).

[8] Malinova, T., Malinov, S., Pantev, N., "Prediction of Surface Hardness After Ferritic Nitrocaburising of Steels using Artificial Neural Networks", Materials Science and Technology, Vol.17, pp. 168-174 (February 2001).

[9] Malinov, S., McKeown, J.J., Sha, W., "Modelling the Correlation between Processing Parameters and Properties in Titanium Alloys using Artificial Neural Network", Computational Materials Science, Vol.21, pp.375-394 (2001).

[10] Calcaterra, S., Campana, G., Tomesani, L., "Prediction of Mechanical Properties in Spheroidal Cast Iron by Neural Networks", Journal of Materials Processing Technology, Vol.104, pp.74-80 (2000).

[11] Lehr, M.A., Rumelhart, D.E., Widrow, B., " The Basic Ideas in Neural Networks", Communications of the ACM, , Vol.37, No.3, pp.87-92, (March 1994).

[12] Beale, M.H., Demuth, H.B., Hagan, M.T., *Neural network Design*, University of Colorado Bookstore, Boulder, CO (2002).

[13] Beale, M., Demuth, H., *Neural Network Toolbox*, *User's Guide Version 4*, The Math Works Inc., Natick, MA (2001).

[14] Foresee, F.D., Hagan, M.T., "Gauss-Newton Approximation to Bayesian Learning", International Conference on Neural Networks, Vol.3, pp. 1930-1935 (1997).

[15] Compos-Velho, H.F., Issamoto, Luz, J.I., Mikki, F.T., Oliviera, P.P.B., "A Neural Network Approach in a Backward Heat Conduction Problem", Proccedings of the IV Brazilian Conference on Neural Networks, pp.019-024 (1999).

[16] Old Dominion University – Department of Electrical and Computer Engineering, "A Comparison of Three Neural Network Architectures for Automatic Speech Recognition ", [online] Sep16, 2002. http://web.odu.edu/engr/speechlab/paper2.pdf

[17] Devore, J.L., *Probability and Statistics for Engineering and the Sciences*, 5th ed., p 339, Duxbury Publishing Co., Pacific Grove, CA (1999).