Predicting compositions and properties of aluminum die casting alloys using artificial neural network

L. Wang*, D. Apelian*, M. Makhlouf*, W. Huang**
*Advanced Casting Research Center Worcester Polytechnic Institute Worcester, MA 01609 USA
**Northwestern Polytechnic University, Xian, China 710072

ABSTRACT

Despite the large number of existing alloys and alloy databases, identifying proper alloys for specific applications still remains a challenge. In order to facilitate the selection and prediction of aluminum die casting alloys and their properties, an electronic database - "i-Select-Al" - has been developed by the Advanced Casting Research Center (ACRC) and the North American Die Casting Association (NADCA). The key to the predictions is the determination of a relationship between alloy properties, chemical composition, and processing variables. Theoretically, these relationships can be "accurately" determined using fundamental physical principles. However, in practice, the underlying mechanisms are not fully understood and difficult to be utilized. In this case, approximate empirical models are considered. In version 1.0 of the software trend equations have been generated. The nature of these trend equations limits the applicability and prediction ability of the software. In order to improve the prediction power; relationships based on an artificial neural network (ANN) were exploited in version 2.0. ANN has proven to be a highly flexible tool, suitable to treat multiple-input conditions and nonlinear phenomena with complex relationships between input and output variables. This article presents the working mechanisms, the programming, and the application of ANN in this project. The results show that ANN is a valuable modeling tool for predicting properties-from-composition and composition-from-properties for aluminum die casting alloys.

RIASSUNTO

Nonostante l'elevato numero di leghe di alluminio disponibili, l'identificazione delle leghe appropriate per specifiche applicazioni rimane un compito arduo. Al fine di facilitare la selezione della lega più appropriata e di predirne le caratteristiche, un database elettronico – "i-Select-Al" – è stato sviluppato congiuntamente dall'Advanced Casting Research Center (ACRC) e dalla North American Die Casting Association (NADCA). Punto nodale della predizione è la determinazione della relazione tra le proprietà della lega d'alluminio, la composizione chimica e le variabili di processo. In teoria, queste relazioni possono essere "accuratamente" determinate partendo dai fondamentali principi fisici. Tuttavia, nella pratica, i meccanismi di funzionamento che stanno alla base della teoria non sono ancora completamente noti e quindi difficili a utilizzarsi. In tal caso si adoperano modelli empirici approssimati. Nella versione 1.0 del software vengono generate le equazioni di "trend". La natura di queste equazioni limita l'applicabilità e la capacità predittiva del software. Al fine di migliorare la capacità predittiva del software, nella versione 2.0 sono state impiegate relazioni basate su reti neuronali artificiali (ANN). Le reti neuronali si sono dimostrate essere strumenti flessibili, adatti al trattamento di sistemi multivariabili e fenomeni non lineari che presentano relazioni complesse tra i dati di ingresso e quelli di uscita. Il presente lavoro illustra i meccanismi di funzionamento, la programmazione e le applicazioni di reti neuronali alla predizione delle proprietà di leghe d'alluminio. I risultati mostrano che le reti neuronali rappresentano validi strumenti di modellazione per la predizione di proprietà-dalla-composizione e composizione-dalle-proprietà di leghe leggere d'alluminio.

KEYWORDS

Property Prediction, Neural Network, Alloy Selection, Aluminum Alloys, Die Casting.

INTRODUCTION

Selecting materials, exploring the potentials of a material, or tailoring the ingredients or properties of a material for given applications are the questions which almost every manufacturing engineer has to face. In order to facilitate material selection, various graphs, tables, electronic databases, expert systems, etc. were developed; however, none of them was

devoted to meet the specific needs of the aluminum die casting industry. Moreover, in most of the existing resources there is only the "selection" ability, i.e., the ability to select an alloy to meet specific property requirements or vice versa. The existing resources lack the "prediction" ability, i.e., the ability to predict the properties from a given chemical composition and vice versa. The prediction ability is very useful and essential in optimizing or tailoring an alloy and for fully utilizing an alloy's potential, and in developing new alloys. In this context, and based on the needs of the die casting industry, ACRC and NADCA initiated a project to develop an electronic tool specific to aluminum die casting alloys that is capable of both alloy selection and prediction. The effort resulted in the software i-Select-Al versions 1.0 and 2.0 [1, 2]. This paper will briefly introduce work done to develop and construct the software with particular emphasis on the version 2.0.

DEVELOPMENT OF AN ELECTRONIC TOOL FOR ALLOY SELECTION AND PROPERTY PREDICTION

The tool was designed so as to have two primary functions: selection and prediction. The selection function allows identifying existing alloys that meet specified requirements. The selection could be performed by inputting the required properties and the software selects the alloy chemistry, or by inputting the alloy composition and software finds a matching (or closest matching) existing alloy(s), and also by inputting the alloy chemistry and the software finds its properties. The prediction function performs the same tasks regardless whether or not the alloy existed. In order to achieve these objectives, the following was needed (1) a comprehensive database to select from, (2) a method, which related the chemical composition to properties and vice versa, and (3) a computer program that can perform these tasks.

DATABASE DEVELOPMENT AND PROGRAM STRUCTURE

The data and information for aluminum die casting alloys can be found in various sources, typically in standards, which are available from different organizations and different countries, and from various handbooks, producers' brochures, research literature, lab reports, patents, computer software, etc. During data mining, we found more than 400 aluminum die casting alloys. presented in many different formats and often the information was inconsistent. Moreover, the available data could have different reliabilities. Some were based on large numbers of measurements using standardized specimens in systematic studies under well-controlled conditions, and others were obtained based on only a few specimens with no information about specimen production and measurement conditions. After analyzing the available information, we discarded some alloys, mostly old, obsolete ones, and ones with uncertainties in the way of measuring their properties, and some that were repetitions. After this elimination process, about 300 alloys remained. These alloys were sorted, formatted consistently, tabulated into groups and used as the database.

For building the interfaces and writing the programs necessary for the electronic tool, two issues needed to be decided upon first. One is the programming language for writing the codes, and the other is the criteria for the selection procedure.

After reviewing several commercially available software packages, Microsoft Access was selected and used for this work. Most personal computers (PCs) have Microsoft Access as part of their standard software package.

The criterion for the selection procedure addresses questions such as how close should an alloy be to the input parameters in order to qualify as potentially meeting the requirements. For example, if a user inputs a few property values as requirements for selecting an alloy, and if the database does not contain an alloy that exactly meets these requirements, but it has

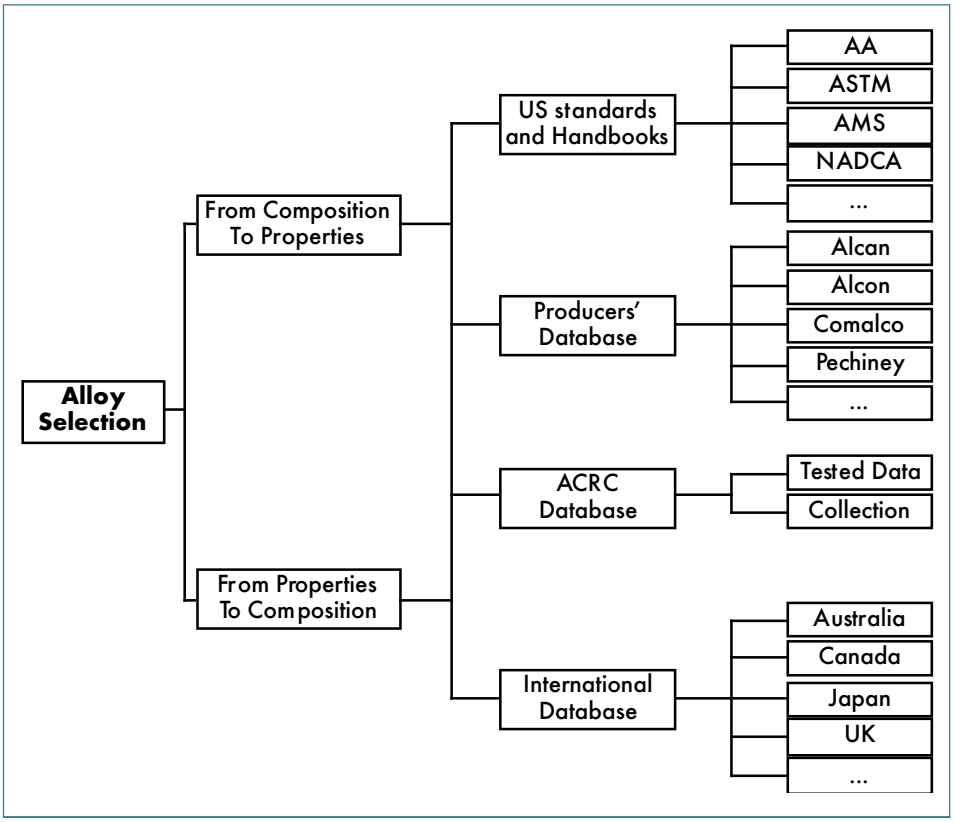

Fig. 1: Flow chart for alloy selection in version 1.0 of i-Select Al.

some alloys that partially meet them. Here, the criteria decide whether these alloys that partially meet the requirement, should be presented as selection or not

The program structure for alloy selection in version 1.0 of the software is shown in Fig. 1 and a typical interface is shown in Fig. 2. All the interfaces were designed to be self-explained and user friendly. The software presents as much information as available and provides various functions to meet the users' needs and the databases were designed to be accessible to the advanced user for adding his/her own data.

The key to alloy prediction is knowing the relationships between chemical composition, processing parameters, and properties. Much of this information exists in the open literature and may be divided into two very general categories: (1) Relationships that are based on physical principles and reflect the essence of the process and the physical and chemical interactions among the factors, and (2) Relationships that are obtained by mathematical means that treat experimentally obtained data as numbers and manipulate this data to obtain relationships between the independent and dependent variables without emphasizing physical meanings. This is the method used in this work.

First, a comprehensive experimental effort was conducted at the Advanced Casting Research Center (ACRC) to relate the chemistry, microstructure, and performance characteristics of aluminum die casting alloys [3]. In this effort, we used a Taguchi design of experiments [4].

Early in the effort, we used multiple regression analysis to derive trend equations that relate alloy chemistry and properties. Multiple regression analysis generates curves that fit the discrete data obtained from experiments to allow estimates at intermediate points. For a given property, this method seeks to derive a single curve (an equation) that represents the general trend of the data. No effort is made to intersect every point, but rather the curve follows the pattern of the points taken as a group. These curves (equations) are used to tell the direction in which an element affects a property, but, in some cases, the predicted values deviate from the measured values.

Later, we used a special analysis of variables method (ANOVA) to derive the trend equations. This was a modified

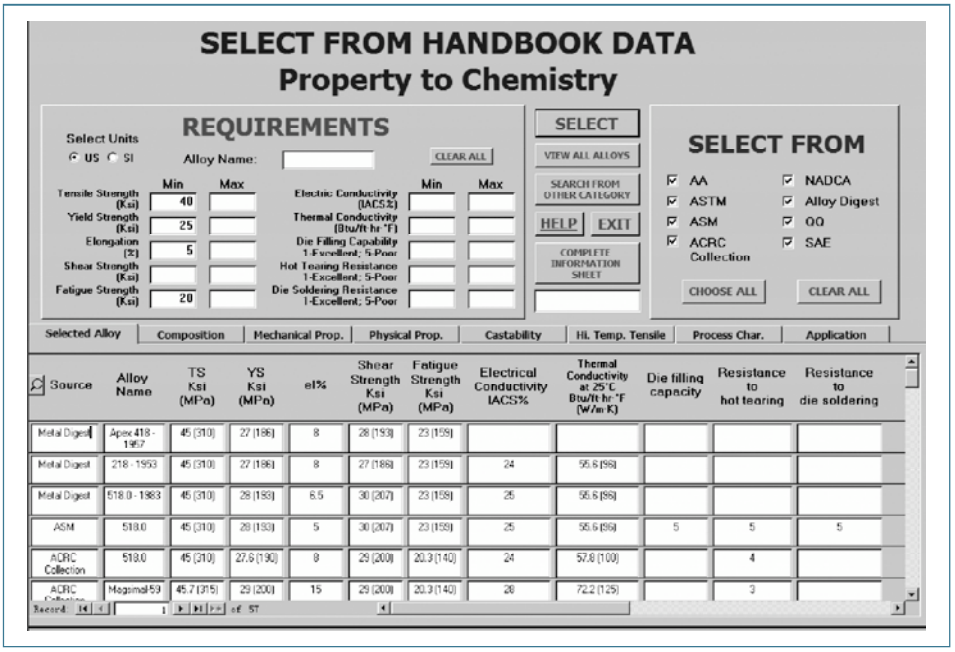

Fig. 2: A typical interface for selecting alloy chemistry from input properties (database obtained from handbooks).

ANOVA designed especially for analyzing data from experiments that were organized using the Taguchi method [4]. The special ANOVA method allows us to derive formulas for optimizing the process parameters. In fact, these formulas are the trend equations and they are used to estimate the results (e.g. properties) when the factors (alloying elements) are within their tested ranges, provided that the effects of the factors are linear and their interactions are negligible. Our experimentally obtained data met these requirements. The elements in the tested composition ranges did not show significant interactions with one another, and for most of the elements the tested ranges were relatively narrow, so assuming that the effects of these elements were linear in the tested ranges did not produce significant errors. These formulas were thus used as the trend equations for the alloy prediction module of i-Select-Al, version 1.0. Trials with the software showed that its predictions are in reasonable agreement with measured values.

DEVELOPMENT OF AN ARTIFICIAL NEURAL NETWORK FOR ALLOY PREDICTION

In order to overcome the limitations of using trend equations and to improve the prediction ability of the software, two different techniques were investigated. These are: (1) the response surface method (RSM), and (2) artificial neural networks. We decided to use artificial neural networks in version 2.0.

Artificial neural networks have been widely used in recent years in many fields and continue to prove to be a highly flexible modeling tool for science and engineering. It is particularly suitable to treat those phenomena that have multiple inputs and have nonlinear and complex relationships between the input and output variables. Attempts, e.g. [5 and 6] have been recently made to use artificial neural networks in predicting the properties of some sand and permanent mold casting aluminum alloys.

The Structure of the Artificial Neural Network – Artificial neural networks are born from developing intelligent systems by simulating the biological structure and function of the human brain. The theory of neural networks is based on neurobiology, mathematics and physics. The basic element of the neural network is the neural cell, the neuron. In general, neural networks are characterized by their architecture, activation functions, and learning algorithms or rules. About 40 neural network models have been proposed and studied. Each type has its

input-output characteristics and applications. For the applications in which both inputs and outputs are numerical data, such as in this work, the most commonly used network is the feed-forward multilayer network with back propagation (BP) training technique. This net obtains its working ability through training the system using the BP learning algorithm. It learns the problems by examples like the human brain. That is, it works based on its previous experiences. The BP learning algorithm is based on the principle of gradually minimizing the errors of the network output by modifying the network parameters. Neurons in the feed-forward neural network are arranged in a layered order. Each neural network is composed of three (or more) different layers: one input, one (or more) hidden and one output layers. Fig. 3 shows a feed-forward network, which has 1 input, 1 output and 1 hidden layer. Each layer consists of a number of neurons (nodes). This kind of network, in which the neuron at a layer receives its inputs only from neurons at the layer in front of it, is called a feed forward neural network [7]. It was this type of the neural network that was used in the present project.

In this network each neuron in the input layer conveys an input value (information) to the network and transmits it to the next layer as value. The output layer is the layer where the output values of the network are generated. The numbers of neurons present at the input and output layers are equal to the input and output variables of the problem, respectively. The hidden layer(s) is placed between input and output layers. The number of hidden layers and the number of neurons in each hidden layer are critical for the convergence rate during training and influence the learning time and accuracy the network can achieve. There are currently no theories or rules for building the hidden layers. The number of layers and the number of neurons in the hidden layers are determined by experience and experimentation with the problem at hand [7].

The Function of the Neurons – The biological neurons show different reactions to different input signals. To model this effect, in the artificial neural network a single number, called weight, is allocated to each input so that each input is multiplied by a weight before being sent to the neuron. The neuron receives the inputs and processes them to generate an output through two functions: a summation function and an activation function, as shown in Fig. 4. In the summation function, the weighted inputs are summed together by simple arithmetic addition to produce the activation, *a* [7].

$$a=\sum_{i=1}^{n} w_i x_i = w_1 x_1 + w_2 x_2 + \ldots w_n x_n \qquad (1)$$

In Eq. (1), w is weight, x is input and i is the index of the input and the weight in the input layer.

The neuron reacts to the activation producing an output and this process is described by the activation function. For the type of neural network used in this project the sigmoid function, $\sigma(a)$, given in Eq. (2), is used as an activation function [7],

$$y=\sigma(a)=\frac{1}{1+e^{-(a-\theta)}} \qquad (2)$$

In Eq. (2), y is the output, a is the activation, and θ is the threshold, and is equal to a at $y=1/2$.

A neuron can receive multiple inputs, but it can generate only one output. The output of the neuron is transmitted along the outgoing connections to serve as input to subsequent neurons.

The Back Propagation Algorithm and Network Training – The artificial neural network is trained using data obtained from experiments. Training is performed in order to find the error by comparing the output value of the network with the target value and then minimizing the difference (error) by modifying the weights. At the start of training, the network creates a set of random weights between the neurons in the successive layers and computes the outputs. The calculated outputs are compared with the target values and the differences are the errors. According to the sign, plus or minus, and magnitude of the error, the program adjusts each weight so as to reduce the total error. Then the calculation is performed once again using the adjusted weights in order to get a new set of outputs, these in turn are compared with the targets, and the weights are adjusted again. This process is repeated until an acceptable total error is reached. At this point, the training is complete with the set of weights

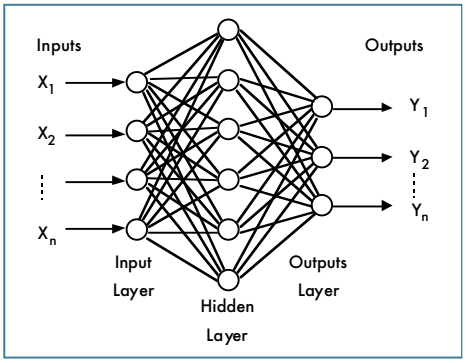

Fig. 3: Structure of a feed forward artificial neural network.

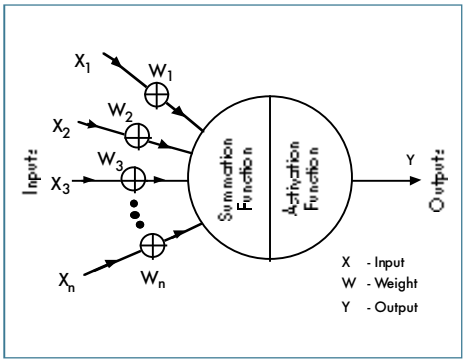

Fig. 4: Function of the neuron.

fixed and the program can be put in use. In use, the neural network can be considered as a black box into which you input the independent variables to get the outputs, the dependent variables, and you do not need to know the physics behind the process.

The Delta Rule – The behavior of the network is completely characterized by the set of weights, so that any function E (error function), which expresses the discrepancy between the desired and the actual network outputs, may be considered to be a function of the weights, i.e., $E=E(w_1,w_2,\ldots,w_{n+1})$, where w_{n+1} is equal to the threshold θ. Therefore, optimizing the weight vector will cause the output to approach the target value, t. The delta rule is the method that is used to optimize the weight vector. It is based on adding a small Δw, which is related to the calculated error, to the weight w used in the last iteration of the error function calculation, and calculating the error again using the modified weight. If the right Δw is chosen, the error should be reduced. The core of

the delta rule is to find the right Δw for the process. This is accomplished as follows.
For a layer with M neurons, for example the output layer, the error E is the sum of the errors e over all neurons.

$$E^p = \sum_{j=1}^{m} e_j = \frac{1}{2}\sum_{j=1}^{M}(t^p_j - y^p_j)^2 \quad (3)$$

and e is defined as:

$$e^p_j = \frac{1}{2}(t^p_j - y^p_j)^2 \quad (4)$$

where p denotes the training pattern, which includes a group of input sets (input pattern) and a group of the corresponding output sets (output pattern), j is the neuron index in the output layer, t is target and y is the output calculated using equation (2) [7].
The Δw is given as:

$$\Delta w_j = \alpha \frac{\delta E}{\delta w_j} \quad (5)$$

In Eq. (5), $\alpha > 0$ is a constant known as the learning rate and governs the learning speed. Based on equations (2) and (5) the following equations can be derived [7].
For the k^{th} weight of the j^{th} neuron in the output layer, Δw_{jk} is given as:

$$\Delta w_{jk} = \alpha\sigma'(a_j)(t^p_j - y^p_j)x^p_{jk} \quad (6)$$

where $\sigma'(a) = d\sigma(a)/da$, $(\sigma'(a)$ is the activation function, defined in equation (2)), and x^p_{jk} is the input from the k th neuron in the hidden layer.
For a neuron in the hidden layer, the target value is unknown, its learning rule (for the kth neuron) can be written as:

$$\Delta w_{ki} = \alpha\sigma'(a_k)\delta_k x^p_{ki} \quad (7)$$

$$\delta_k = \sum_{j \in I_k} \delta_j w_{jk} \quad (8)$$

where $\delta_j = (t^p_j - y^p_j)$ and I_k is the set of neurons that take an input from the hidden neuron k.
In a fully connected network such as the one used in this work, I_k is the whole of the output layer [7].
The learning speed is governed by the learning rate a. A small a will guarantee stability of the network training but is rather inefficient. If a is increased too much, learning will become unstable; the net will oscillate back and forth across the error minimum, or it may wander aimlessly around the error landscape. In some cases where a is too large, learning is "bumpy" with many undulations. In order to improve the stability of the network training, the momentum term $\lambda\Delta w(n-1)$ is introduced into the learning rule. It is the momentum constant $\lambda(0<\lambda<1)$ multiplied by the previous weight change [7].
The computer program in this project was written using a back propagation training algorithm and the above delta training rules. The program considered only alloy chemistry and alloy properties and did not consider other factors such as casting conditions, cast part dimensions, etc. Accordingly, only data that was obtained using similar measurement procedures on standardized test specimens that were produced by similar casting conditions was used in training the software. Relevant data on 40 alloys was used for training the neural network. The data was obtained at ACRC by measurements of properties of alloys with precisely controlled compositions and also from some standard alloys.

RESULTS AND DISCUSSION

Figs. 5, 6, and 7 show comparisons between computer-predicted and measured yield strength, elongation, and thermal conductivity for 25 different alloys. In these figures, alloys 1 to 24 were used in training the neural network and alloy 25 was used to gage the accuracy of the computer prediction. Similarly, Fig 8 shows comparisons between the computer-predicted and the target chemistry of 2 alloys. One alloy, alloy 8, was used in training the neural network and the second alloy, alloy MPI380, was used for gagging the accuracy of the computer prediction.
We notice that the accuracy is a little lower when predicting alloy chemistry from alloy properties than when predicting alloy properties from alloy chemistry. This is probably related to the ratio of inputs to outputs used in the training for each case. The ratio of inputs to outputs is 10 to 5 in the later, and 5 to 10 in the former. The results in Figs 5, 6, 7, and 8 show that the errors between the output and the target values for the training data sets are very small.

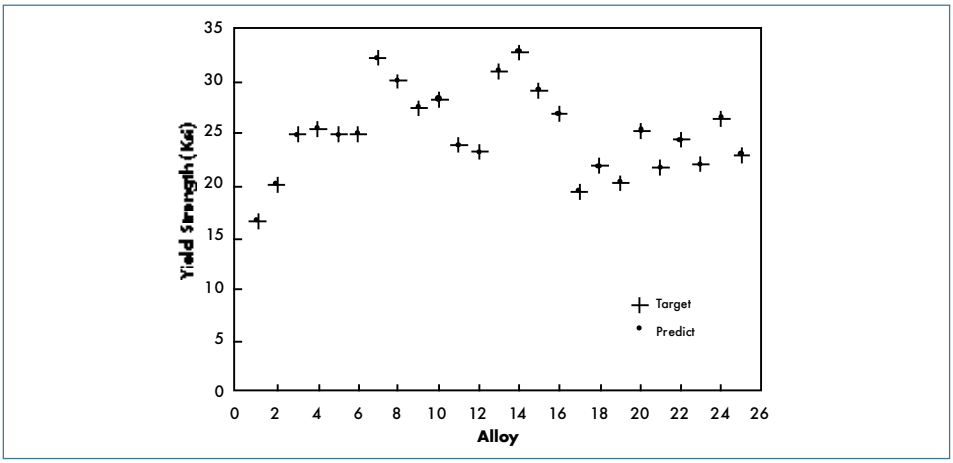

Fig. 5 Predicted and measured yield strength of 25 alloys. Alloys 1 to 24 were used in training the network, and alloy 25 was used for gagging the accuracy of the computer prediction.

SUMMARY AND CONCLUSIONS

The artificial neural network is a powerful tool for treating phenomena that have multiple inputs (independent variables) and have nonlinear and complex relationships between the input and output variables. In predicting the composition and properties of aluminum die casting alloys, the artificial neural network gives very small errors between the input and target values for the training sets of data and has good accuracy in predicting alloy properties from alloy chemistry. The accuracy in

predicting alloy chemistry corresponding to a specified set of alloy properties is not as good, but is reasonable and useful. The success of the artificial neural network depends to a very large extent on the database that is used in training it. The accuracy, quantity of entries, and coverage of the database determine the accuracy and coverage of the artificial neural network. The accuracy and coverage of this artificial neural network can certainly be improved by enriching the database. While this effort considered only the prediction between alloy chemistry and properties, it is possible to incorporate other factors such as casting conditions, cast part dimensions, etc. into the network.

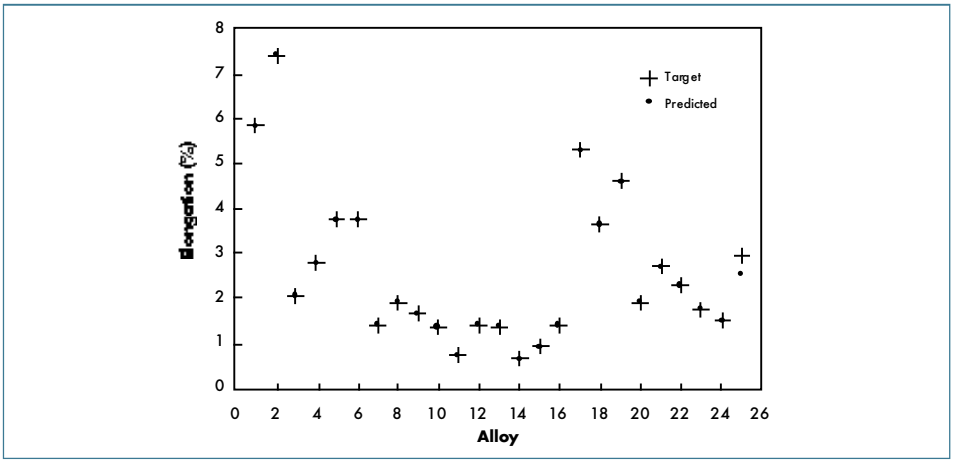

Fig. 6 Predicted and measured elongation of 25 alloys. Alloys 1 to 24 were used in training the network, and alloy 25 was used for gagging the accuracy of the computer prediction.

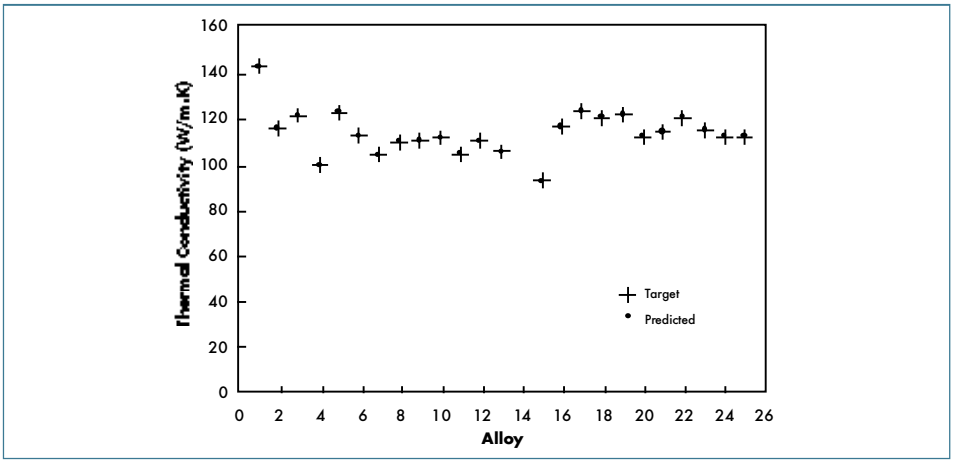

Fig. 7 Predicted and measured thermal conductivity of 25 alloys. Alloys 1 to 24 were used in training the network, and alloy 25 was used for gagging the accuracy of the computer prediction.

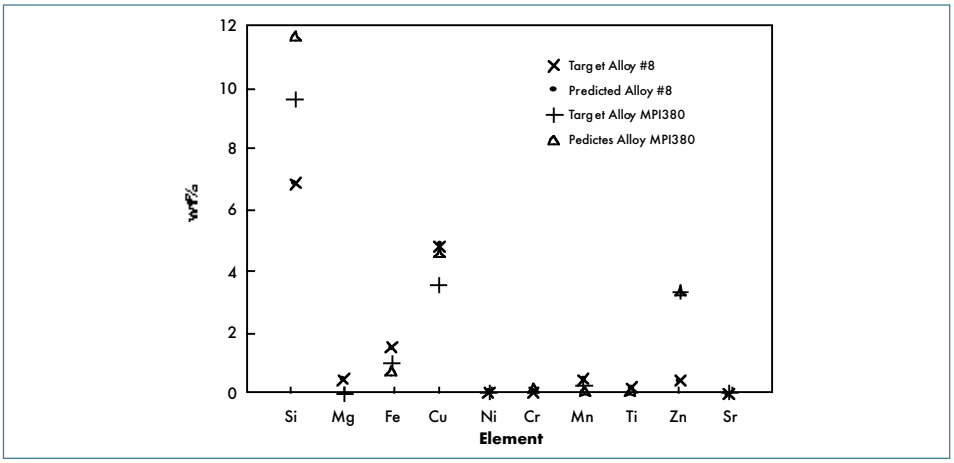

Fig. 8 Predicted and measured compositions of alloys 8 and MPI380. Alloy 8 was used for training the network and alloy AMC380 was used for gagging the accuracy of the computer prediction.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the continued support of the North American Die Casting Association (NADCA), the American Metal Casting Consortium (AMC), and the member companies of the Advanced Casting Research Center (ACRC).

REFERENCES

[1] "i-Select-Al Full Version", NADCA Publications, Item #215-CD.

[2] M.M. Makhlouf, D. Apelian and L. Wang, "i-Select-Al: an Electronic Tool for the Selection and Design of Aluminum Die Casting Alloys", Die Casting Engineer, Vol. 48, No. 2, March 2004, pp 26-31.

[3] M. M. Makhlouf, D. Apelian, and L. Wang, Microstructure and Properties of Aluminum Die Casting Alloys, NADCA, 1999.

[4] G. Taguchi, System of Experimental Design, UNIPUS, Kraus International Publication, 1987.

[5] D. Emadi and L.-A. Sullivan, "Application of Artificial Neural Networks in Prediction of Mechanical Properties of Cast Aluminum Alloys", AFS Transactions 2004, paper 042.

[6] D. Emadi, M. Sahoo, T.Castles, and H. Alighanbari, "Prediction of Mechanical Properties of As-cast and Heat-Treated Automotive Al Alloys Using Artificial Neural Networks", TMS Light Metals 2001, pp 1069-76.

[7] Kevin Gurney, An Introduction to Neural Networks, UCL Press, 1997.