Computational Materials Science 44 (2009) 1231–1235

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Mechanical behavior modeling of nanocrystalline NiAl compound by a feed-forward back-propagation multi-layer perceptron ANN

M. Yazdanmehr [a,*], S.H. Mousavi Anijdan [b], A. Samadi [c], A. Bahrami [b]

[a] Physics Department, Isfahan University, Isfahan, Iran
[b] Materials Science and Engineering Department, Sharif University of Technology, Tehran, Iran
[c] Materials Science and Engineering Department, Sahand University of Technology, Tabriz, Iran

ARTICLE INFO

Article history:
Received 24 July 2008
Received in revised form 12 August 2008
Accepted 16 August 2008
Available online 5 October 2008

PACS:
61.46.Df
61.46.Hk
61.82.Rx
62.25.+g

Keywords:
Nanocrystalline compounds
NiAl
Intermetallic
Mechanical properties
Neural network

ABSTRACT

In this paper, an artificial neural network (ANN) model has been developed to predict the yield and tensile strengths of hot pressed NiAl intermetallic compound based on the experimental data from Albiter et al. [A. Albiter, M. Salazar, E. Bedolla, R.A.L. Drew, R. Perez, Mater. Sci. Eng. A 347 (2003) 154]. The predicted results, with a correlation relation between 0.9791 and 0.9921, show a very good agreement with the experimental values. Furthermore, the sensitivity analysis was performed to investigate the importance of the effects of chemical composition and temperature on the mechanical behavior of hot pressed NiAl intermetallic compound.

1. Introduction

The Al–Ni system has been extensively studied because of its excellent properties such as lower density, excellent thermal conductivity, and good oxidation resistance [1]. Among different alloys in this system, NiAl is an important intermetallic compound for high temperature applications including coatings and metal matrix composites (MMCs). In addition, there is a great potential to replace the conventional Ni-based superalloys with the NiAl intermetallic compound [2].

The mechanical properties of NiAl intermetallic compound can be enhanced by the addition of some microalloying elements and reduction of grain size [3–6]. In addition, the consolidation technique is effective in final mechanical properties. Although the spark plasma sintering is the modest technique in order to consolidate the nanocrystalline NiAl powders [7], the hot isostatic pressing technique, low energy reaction synthesis and rapid solidification routes have also been used for the consolidation [8–10].

The mechanical properties of NiAl intermetallic compound are a complex function of chemical composition and fabrication techniques. It has been demonstrated that the artificial neural networks (ANNs) are capable to model the mechanical properties in complex systems with reasonable accuracy [11–14]. At present, ANNs are gaining tremendous momentum for prediction purposes in the field of nanomaterials because of their specific features such as non-linearity, adaptivity (i.e. learning from inputs parameters), generalization, and model independence (a priori model is not need).

The aims of this paper are (1) to examine the applicability of the ANN in prediction of mechanical properties of nanocrystalline NiAl intermetallic compounds from the available data, and (2) to determine the influential parameters on the mechanical properties via sensitivity analysis.

2. Basics of ANNs

The ANNs are mathematical modeling tools that are particularly useful in the field of prediction and forecasting in complex settings. Historically, there were meant to operate through simulating, at a

* Tel.: +98 21 66887791.
E-mail address: yazdanmehrmaryam@yahoo.com (M. Yazdanmehr).

0927-0256/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.commatsci.2008.08.006

simplified level, the activity of the human brain. The ANN accomplishes this through a large number of highly interconnected processing elements (neurons), working in unison to solve specific problems, such as forecasting and pattern recognition. Each neuron is connected to certain of its neighbors with varying coefficients or weights that represent the relative influence of the different neuron inputs to other neurons [12].

There are many kinds of ANNs. One of the most common types of neural networks is the feed-forward, where the information is transmitted in a forward direction only. This is the type employed in this work. In this study, the tan-sigmoid activation function (Eq. (1)) is used for the input and hidden layers, and also is used as activation function between hidden layers and the output

$$f(x) = \tan\left(\frac{1}{1+\exp(-x)}\right) \quad (1)$$

The tan-sigmoid activation functions for the input and hidden neurons and also for the hidden and output neuron are needed to introduce non-linearity into the network. Without non-linearity, hidden layers would not make nets more powerful than just plain perceptrons (which do not have any hidden units, just input and output units). Sigmoid activation functions are usually preferable to threshold activation functions. Networks with threshold functions are difficult to train because the error function is stepwise constant, hence the gradient either does not exist or is zero, making it impossible to use back-propagation or more efficient gradient-based training methods. With sigmoid function, a small change in the weights will usually produce a change in the outputs, which makes it possible to assess the changes in weights. With threshold function, a small change in the weights will often produce no change in the outputs [15].

Network training is a process by which the connection weights and biases of the ANN are adapted through a continuous process of simulation by the environment in which the network is embedded. The primary goal of training is to minimize an error function by searching for a set of connection strengths and biases that causes the ANN to produce outputs that are equal or close to targets. The minimization procedure relies on a numerical optimization of a non-linear objective function. A number of optimization routines can be used. In practice, the Levenberg–Marquardt routine often finds better optima for a variety of problems than do the other optimization methods [16]. The learning process may be supervised or unsupervised. Supervised learning incorporates an external teacher, so that each output neuron is told what its desired response to input signals ought to be. The learning process, global information may be required. An important issue concerning supervised learning is the problem of error convergence, i.e. the minimization of error between the desired and computed unit values. So, the main purpose will be the determination of a setting of weights which minimizes the error.

Unsupervised learning, on the other hand, uses no external teacher and is based upon only local information. In fact, for most varieties of unsupervised learning, the targets are the same as the inputs [15]. It is also referred to as self-organization, in the sense that it self-organizes data presented to the network and detects their emergent collective properties. In other words, unsupervised learning usually performs the same task as an auto-associative network, compressing the information from the inputs. Unsupervised learning is very useful for data visualization. In this work, supervised learning was used.

The architecture of a neural network consists of a description of the network's number of layers, the number of neurons in each layer, each layer's activation function, and how the layers connect to each other [16–18]. Learning rate parameter may also play an important role in the convergence of the network, depending on the application and the network architecture [19]. The learning rate is used to increase the chance of avoiding the training process from being trapped in a local minimum instead of the global minimum.

The network architecture and learning rate were determined in this work using a trial-and-error approach, as explained in the later discussion.

Table 1
The range of the processing conditions used as input variables in the ANN model

Variable	Possible values
Fe, Ga, Mo	2 and 6 (at.%)
Temperature	1200 °C (30 min) and 1500 °C (15 min)

3. Materials and data collection

The initial data used for this study have been selected from a published work by Albiter et al. in which the mechanical properties of a nanocrystalline NiAl intermetallic alloy with Fe, Ga and Mo additions under two different hot pressing conditions have been measured [2]. In this investigation 33 samples have been produced by hot pressing. Among these data, 27 samples were selected for the ANN training process, while the remaining six samples were used to verify the generalization capability of ANN. Table 1 shows the high and low levels of effective variables on mechanical properties of nanocrystalline NiAl intermetallic compound.

4. Results and discussion

4.1. Model calibration (training)

The models were calibrated using the back-propagation algorithm, as it has already been used successfully for the prediction of the mechanical properties and other materials variables and, unlike second order optimization algorithms, has the ability to escape local minima in the error surface. Optimal values of the parameters

Table 2
Characterization of the best network

ANN model	Geometry (I–H–O)	Learning rate	Momentum rate
1	7–15–2	0.4	0.2

I: the number of model inputs.
H: the number of nodes in hidden layer.
O: the number of model outputs.

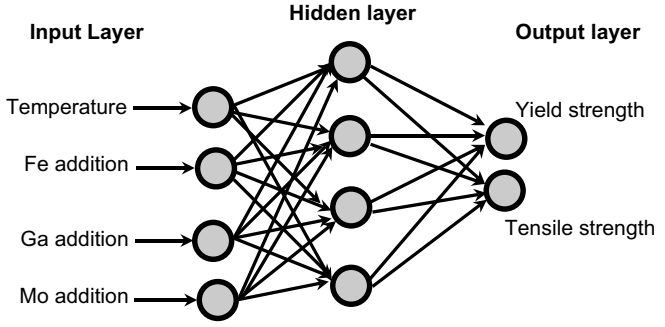

Fig. 1. Schematic diagram of multi-layer neural network which has been used in this investigation.