Application of artificial neural network to predict flow stress of as quenched A357 alloy

X. W. Yang*, J. C. Zhu, Z. H. Lai, Y. R. Kong, R. D. Zhao and D. He

In order to establish an accurate thermal stress mathematical model of the quenching operation for A357 alloy, isothermal compression tests were performed in the temperature range of 200–500°C and the strain rate range of 0·001–1 s^{-1} on as quenched A357 (Al–7Si–0·6Mg) alloy. Based on the experimental results, the deformation behaviour of as quenched A357 alloy was investigated in terms of artificial neural network with a back propagation learning algorithm. Using the deformation temperature, strain rate and strain were used as inputs, and flow stress was used as the output in the network. The average absolute relative error between the predicted results that used the back propagation network and the experimental data obtained from compression tests is 2·89%, which indicates that this artificial neural network model is able to predict the flow stress with high precision. Therefore, it can be used as an accurate thermal stress model to solve the problems of quench distortion of parts.

Keywords: A357 alloy, Artificial neural network, Flow stress

Introduction

Al–7Si–0·6Mg (A357) alloy has been widely applied in aerospace, automotive industries and other aspects because of its excellent castability, good corrosion resistance and high strength/weight ratio in the heat treated condition.[1–3] The thin walled workpieces (e.g. frame, thin walled beam and wall panel) have been casted using this alloy, which has a wide range of applications in the aerospace industries. The thin walled workpieces are routinely heat treated to the T6 state in order to develop an adequate mechanical property. The T6 heat treatments include solution heat treatment, quench and age hardening. Stresses and strains caused by quenching treatment may lead to distortion of the thin walled workpieces. In order to calculate and analyse the residual stress of thin walled workpieces during quenching treatment, an accurate thermomechanical model should be founded. Therefore, the constitutive behaviour of A357 is needed to be established during the quenching operation.

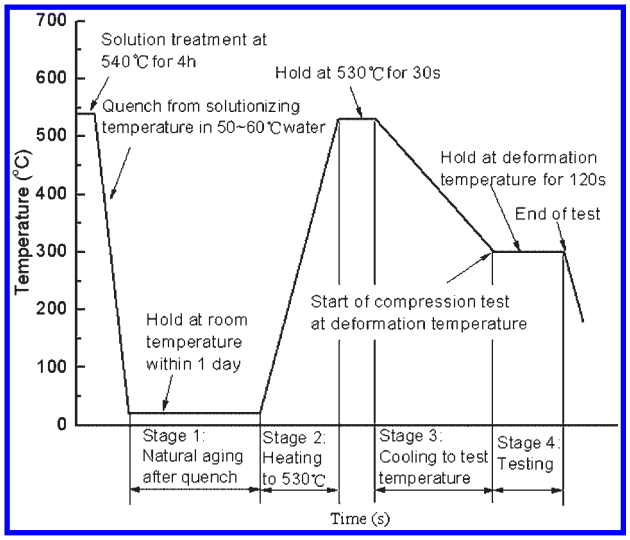
The constitutive behaviour of materials is usually expressed by a non-linear equation that is derived from the relationship of flow stress, strain, strain rate and temperature. In the 1960s, Sellars and Tegart proposed the hyperbolic sine law to characterise the flow stress, and then a large amount of work had been performed in the past several decades to describe the relationship between the flow stress and the process parameters in the constitutive models.[4–8] These constitutive models are either empirical or semiempirical in nature. The empirical or semiempirical method requires the selection of a mathematical model in which the values of the constants are determined using multivariate non-linear regression. Although these constitutive equations attempt to describe non-linear relations among flow stress, strain, strain rate and temperature, these are usually restricted to regimes where specific deformation mechanisms generate. It is difficult to use any simple constitutive model to describe the complicated correlation of flow stress with their affecting factors (microstructure and processing parameters). In addition, when new experimental data are added, regression parameters must be recalculated. In order to expand the study domains, the different equation parameters and/or several independent equations need to be used.

Recently, artificial neural network (ANN) has been widely applied in describing the constitutive behaviour of materials.[9–13] The ANN method is unlike the empirical model, however, which is capable of treating non-linear problems and complex relationships. As far as the constitutive behaviour of materials is concerned, the approach could be able to predict the flow stress through self-organisation without taking the deformation mechanisms into account. Generally, good self-organising capacity of the ANN network is obtained from a lot of experimental data. Therefore, a series of isothermal compression tests were carried out on samples of the A357 alloy in quenching state at the strain rate range of 0·001–1 s^{-1} and temperature range of 200–500°C. Therefore, on the basis of a lot of experimental data based on compression tests, an ANN with a back propagation (BP) learning algorithm was proposed, which was about the as quenched A357 alloy.

School of Materials Science and Engineering, Harbin Institute of Technology, Harbin 150001, China

*Corresponding author, email hgdrali@sina.com

Published by Maney on behalf of the Institute
Received 11 October 2010; accepted 22 December 2010
DOI 10.1179/1743284710Y.0000000051



1 Thermal history used in compression tests (from Ref. 16)

Experimental

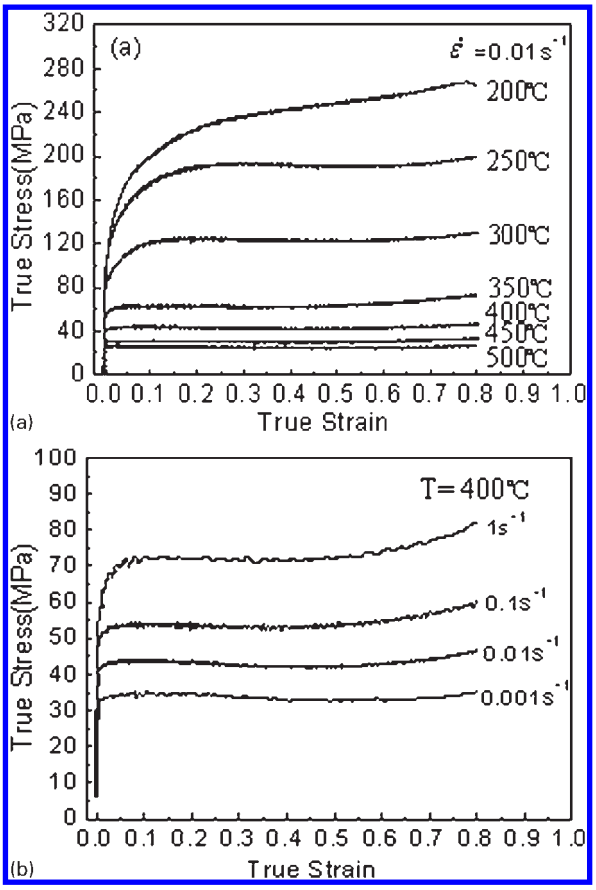
The main chemical composition of commercial A357 aluminium alloy used in the present investigation is as follows: Al–6·83Si–0·51Mg–0·18Ti–0·04Cu–0·03Fe–0·02Be (wt-%). Cylindrical samples for compression tests were machined according to ASTM E9 standard specification[14,15] with a diameter of 8 mm and a height of 12 mm.

Estey *et al.*[16] depicted the thermal history used in the compression tests for A356 alloy, which was reproduced in Fig. 1. The A356 and A357 alloys belong to the Al–Si–Mg cast alloys and contain 7%Si. Therefore, in the present paper, the authors conducted the tests for A357 alloy similar to the thermal compression tests for A356 alloy.[16] According to the literature,[16] before thermal compression tests, the samples were solutionised at 540°C for 4 h in a resistance furnace and quenched in water at 50–60°C. Then, the samples were maintained at room temperature. Compression tests were conducted within 1 day after solution treatment. Hot compression tests were carried out in a Gleeble-1500D thermal simulator at the strain rate of 0·001–1 s^{-1} and temperature range of 200–500°C until the engineering compressive strain reached 60%. The lubricants of graphite mixed with machine oil were adopted so that friction at the specimen/die could be reduced. In order to offset any precipitation that may have occurred due to natural aging, the samples were heated up to 530°C at a heating rate of 10°C s^{-1}, kept for 30 s to solutionise, cooled to the temperature required in the compression test at a cooling rate of 5°C s^{-1}, then kept for 120 s for structural homogeneity and finally compressed by the Gleeble-1500D thermal simulator.

True stress–strain curves of the as quenched A357 alloy can be easily plotted from the data recorded by the thermal simulator during hot compression tests. Figure 2 shows the typical true stress–strain curves of the as quenched A357 alloy measured from deformation tests.

Establishment of ANN model

In the present study, the ANN with a BP learning algorithm was established to predict the flow stress of the as quenched A357 alloy during hot compression. The ANN model comprises an input layer, one or two hidden



a at different deformation temperatures; *b* at different strain rates

2 Typical true stress–strain curves of as quenched A357 alloy

layers and an output layer. Deformation temperature, strain and strain rate were used as the input variables, while flow stress was taken as the output.

To set up the present BP ANN, the data from compression tests were divided into two subsets: a training set and a testing set. In the present work, the training set accounts for ~75% of the original dataset, while the remaining dataset was left for the testing set. In the present paper, the input and output values have a large difference, so the authors have to standardise them in the range between [0·1, 0·9] using the following functions

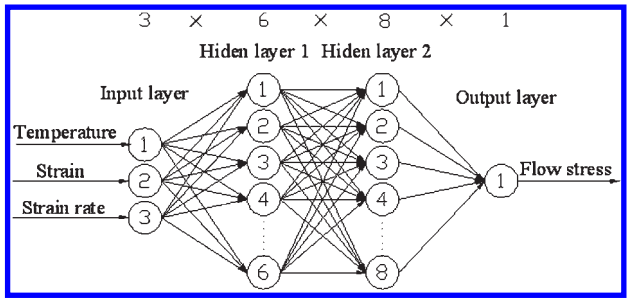
$$x_i = 0{\cdot}1 + 0{\cdot}8\left(\frac{x - x_{\min}}{x_{\max} - x_{\min}}\right) \quad (1)$$

where x are the primary data from compression tests and x_i, $x_{\max}$ and $x_{\min}$ are the normalised, maximum and minimum values of x respectively. When the high precision of network is attained, the data of trained and tested results converted to their original form, which can be indicated as

$$x = \frac{(x_i - 0{\cdot}1)(x_{\max} - x_{\min})}{0{\cdot}8} + x_{\min} \quad (2)$$

In the output layer of the BP model, a linear function was selected as the transfer function, while the transfer functions in the hidden layers obeyed sigmoidal functions (tan sigmoid and log sigmoid), which are defined as

$$f(x) = \tan\mathrm{sig}(x) = \frac{2}{1 + \exp(-2x)} - 1 \quad (3)$$



3 Schematic of BP ANN

$$f(x) = \log sig(x) = \frac{1}{1+\exp(-x)} - 1 \quad (4)$$

The performance criterion for the training of the BP network is commonly evaluated by the mean square error (MSE), which tries to minimise the average squared error between the desired and actual output values over all the examples pairs, and the function of the MSE can be shown as

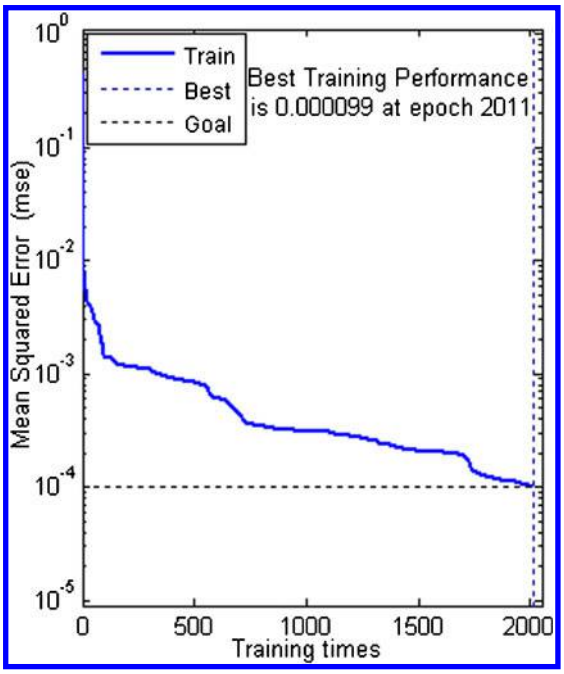
$$MSE = \frac{1}{N}\sum_{i=1}^{N} e(i)^2 = \frac{1}{N}\sum_{i=1}^{N} [d(i) - a(i)]^2 \quad (5)$$

where $d(i)$ is the desired output value from the true stress–strain curves, $a(i)$ is the actual output value from the ANN and N is number of desired data in the study.

One important issue with an ANN is to decide how to train the network. The authors are concerned with two different styles of network training: sequential training and batch training. In sequential training, the weights and biases of the network are updated each time an input is presented to the network. Batch training is when the weights and biases are only updated after all the inputs and targets are presented. For sequential training, the order in which the input vectors appear is important. However, for batch training, the order is not important, and if we had a number of networks running in parallel, we could present one input vector to each of the networks. According to the above description of the two different styles of network training, we have used a batch mode of training in the present paper.

The good performance of the network includes two aspects: the acceptable value of error between the ideal output and the actual output and the stable performance of network. In the present study, for an ANN structure with 1 (input layer)–1 (hidden layer)–1 (output layer), the average absolute relative error between the ideal output and the actual output is ~8%, but for an ANN structure with 1 (input layer)–2 (hidden layer)–1 (output layer), the error is only ~3%. In addition, the performance of the network with hidden layer 2 is stable than the one with hidden layer 1. Therefore, the ANN structure with 1 (input layer)–2 (hidden layer)–1 (output layer) was used in the present study to predict the flow stress of the as quenched A357 alloy.

There is no theoretical method to determine the values of parameters of the ANN model. According to traditional research, the values of these parameters were selected by trial and error, which was used in the present paper. In the traditional ANN literature,[17–20] the learning rate and the momentum constants are usually considered between 0 and 1. An appropriate selection of the learning rate is very important to train ANN. The



4 Curve of training error

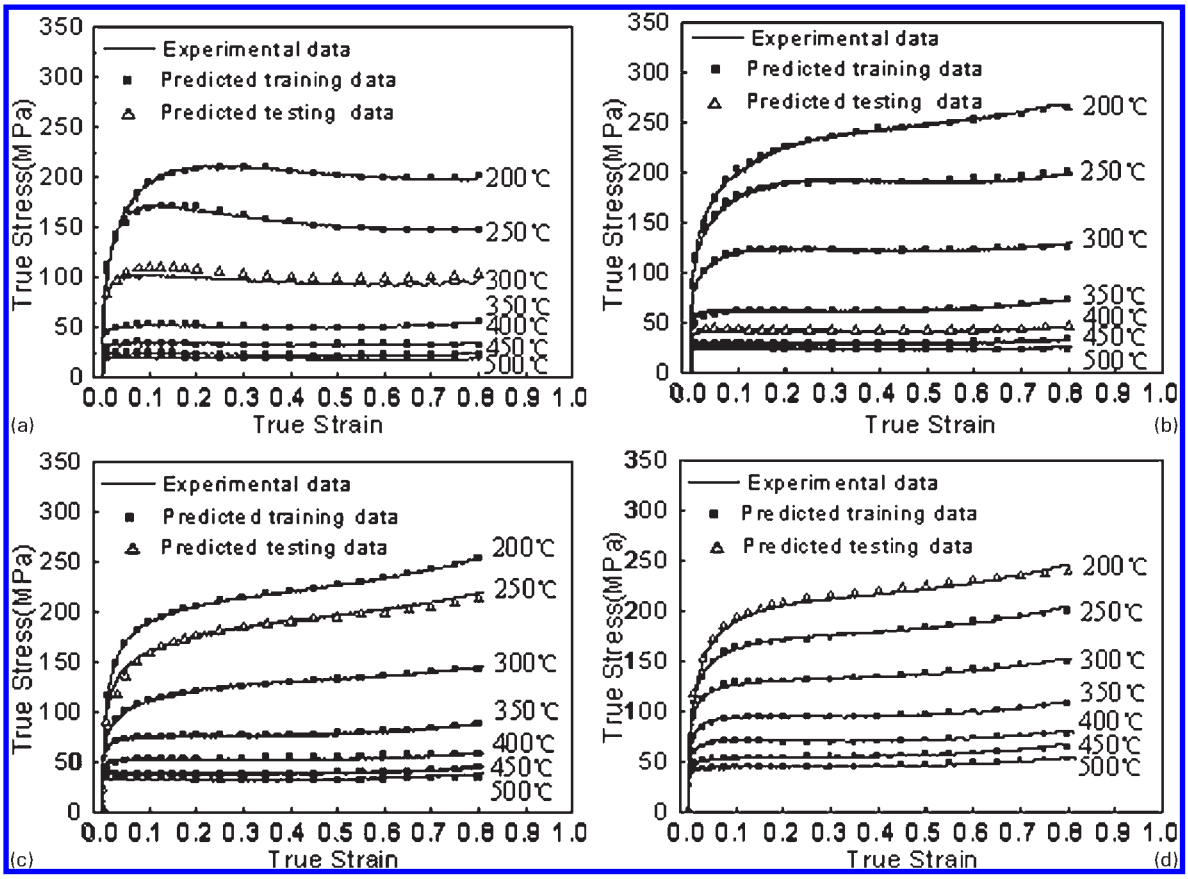
value of the learning rate should not be too high. A too high learning rate may result in an unstable state. On the other hand, if the value of the learning rate is too low, then it may take too long to complete the course of training the network. Therefore, an appropriate learning rate not only ensures the stability of the network but also guarantees an acceptable training time of the network. In the present case, we choose the value of 0·1 for the learning rate. A momentum factor is used to improve the speed of network training. It takes into account the extent to which a particular weight was changed on the previous iteration. When the momentum factor is equal to 0, the change in weights is obtained by the gradient descent approach. When the momentum factor is equal to 1, the change in weights is set to equal to the change in the previous iteration, and the part of the change generated by the gradient descent approach will be ignored. For this reason, a higher value of momentum factor is beneficial to find the set of weights that provides the best performance of the network. In the present case, the authors choose the value of 0·60 for the momentum factor.

After training, a momentum rate of 0·6, a learning rate of 0·1, an MSE of 0·0001 and epochs of 10 000 were fixed for the model. In addition, six neurons were used in hidden layer 1, and eight neurons were used in hidden layer 2. The structure of the ANN model is shown in Fig. 3.

Results and discussion

Using the training set, the neural network had been trained by the learning algorithm. Figure 4 presents the result of MSE for training. It can be seen that the training lasted until the error goal was reached and the training terminated at 2011 epochs. Each epoch is a step that passes through inputs, hides and outputs in the training process of anANN. Here, the value of goal error was set as 0·0001, and the value of MSE for network training reached 0·000099 at epoch 2011.

When the MSE of network training achieved the prescribed goal, the testing set can be used for network



a strain rate=0·001 s^{-1}; *b* strain rate=0·01 s^{-1}; *c* strain rate=0·1 s^{-1}; *d* strain rate=1 s^{-1}

5 Comparison of ANN results with experimental data

testing. In the present paper, the correlation coefficient R, the average absolute relative error e_{abr} and the average absolute error e_{ab} were used to assess the performance of the training and testing network, and the equations of R, e_{abr} and e_{ab} can be described as

$$R=\frac{\sum_{i=1}^{N}(E_i-\bar{E})(P_i-\bar{P})}{\left[\sum_{i=1}^{N}(E_i-\bar{E})^2\sum_{i=1}^{N}(P_i-\bar{P})^2\right]^{1/2}} \quad (6)$$

$$e_{abr}=\frac{1}{N}\sum_{i=1}^{N}\left|\frac{E_i-P_i}{E_i}\right|100\% \quad (7)$$

$$e_{ab}=\frac{1}{N}\sum_{i=1}^{N}|E_i-P_i| \quad (8)$$

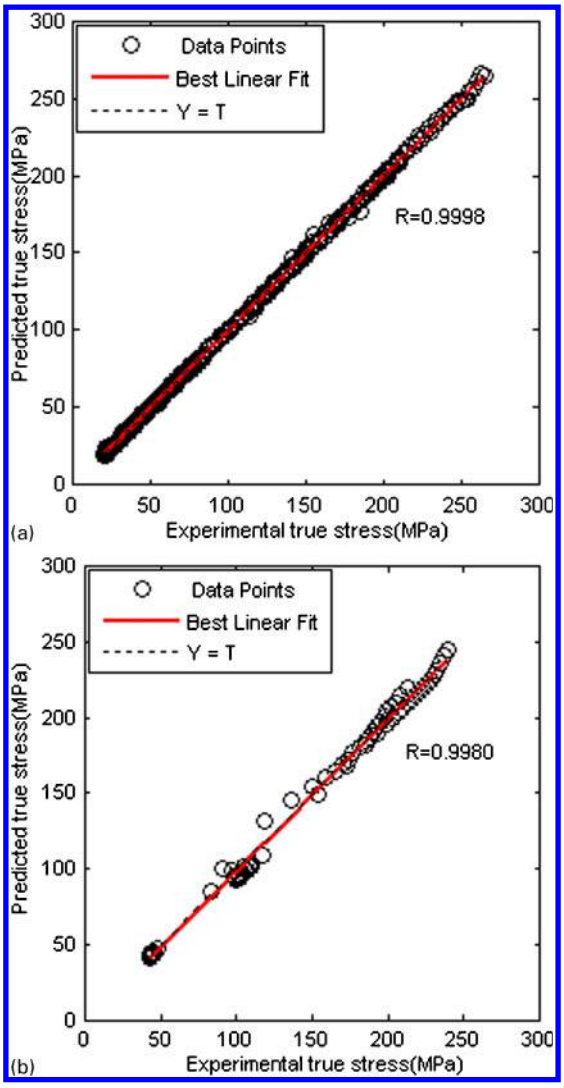
where E is the experimental data obtained from the true stress–strain curves, P is the predicted value from the BP model and $\bar{P}$ and $\bar{E}$ are the mean values of E and P respectively. N is the total number of desired data in the present study.

The testing dataset is composed of four sets of data extracted from four groups of experimental stress–strain curves. Figure 5 shows the comparison of ANN results with experimental data in different strain rates. As shown in Fig. 5, the ANN model presents a higher accuracy in predicting true stress. The performance results (e_{abr} and e_{ab}) of the BP network model are provided in Table 1. It can be seen that average absolute relative errors e_{abr} between the predicted and experimental flow stresses were 1·07 and 2·89% for the training and testing datasets respectively. In addition, it can be found that the value of average absolute error e_{ab} between the predicted data and the experimental data for training was 0·85 MPa while that of the testing dataset was 3·40 MPa. Therefore, the authors can conclude that the generalisation performance of the BP network should be satisfactory.

Figure 5 indicates that the increase in strain rate hardly influences the flow stress of the as quenched A357 alloy at low temperatures ($\leqslant$350°C), while in the experiments carried out above 350°C, the flow stress values were found to be dependent on temperature and strain rate (decreasing with increasing deformation temperature and/or increasing with increasing strain rate). For the three higher strain rate experiments (0·01, 0·1 and 1 s^{-1}), the flow stress decreases with increasing strain rate at low temperatures ($\leqslant$250°C), as shown in Fig. 5*b–d*. For the low strain rate experiments (0·001 s^{-1}), the behaviour is more complex compared with the three higher strain rates (0·01, 0·1 and 1 s^{-1}). As shown in Fig. 5*a*, at low temperatures ($\leqslant$300°C), the material shows a softening behaviour when the strain is greater than a certain value. The 250°C experiment shows a softening behaviour for strains greater than

Table 1 Error of ANN model for training and testing prediction

	True stress prediction	
Error	Training dataset	Testing dataset
e_{abr}/%	1·07	2·89
e_{ab}/MPa	0·85	3·40



a training result; *b* testing result

6 Correlation between predicted data from neural network and experimental data

~0·12. It is also found that the softening behaviour in 200 and 300°C experiments is weaker than the one in the 250°C experiment.

In order to investigate the network capability thoroughly, the regression analysis between the predicted data and the experimental data should be adopted, and the correlation coefficient should be calculated. The correlation coefficient between the ANN predicted flow stresses and the actual experimental test results for both training and testing sets is given in Fig. 6. It is seen that the degree of best linear fit line is very close to the 45° line, revealing close agreement between predicted data and experimental flow stresses. In the meantime, the correlation coefficient was found to be 0·9998 and 0·9980 for both training and testing data. The results show that the correlation coefficient was positive and close to 1, indicating that there is a strong positive linear correlation relationship between predicted data and experimental data. Thus, when the values of experimental data increase, the predicted data values will also increase. Therefore, it can be concluded that the BP model proposed in the present paper is suitable to predict the flow stress of the as quenched A357 alloy, which has a higher precision.

Conclusions

In the present paper, an ANN model with the BP learning algorithm has been established to predict the flow stress of the as quenched A357 alloy. The results show that the mean relative error between the predicted flow stress of the ANN model and the experimental data is 2·89%, which demonstrates that the predicted ANN model has much higher precision and less errors and the predicted values accord with the experiment data basically. Therefore, the BP ANN model proposed in the present paper can be used as the accurate predicting model of the constitutive relations of the as quenched A357 alloy.

The material behaviour observed in the deformation tests is complex. For the four strain rate experiments (0·001, 0·01, 0·1 and 1 s^{-1}), the increase in strain rate hardly influences the flow stress of the as quenched A357 alloy at low temperatures ($\leqslant$350°C), while in the experiments carried out above 350°C, the flow stress values decrease with increasing deformation temperature and/or increase with increasing strain rate. For the three higher strain rate experiments (0·01, 0·1 and 1 s^{-1}), the flow stress decreases with increasing strain rate at low temperatures ($\leqslant$250°C). For the low strain rate experiments (0·001 s^{-1}), at low temperatures ($\leqslant$300°C), the material shows a softening behaviour when the strain is greater than a certain value. The softening behaviour in the 200 and 300°C experiments is weaker than the one in the 250°C experiment.

References

1. O. S. Es-Said, D. Lee, W. D. Pfost, D. L. Thompson, M. Patterson, J. Foyos and R. Marloth: *Eng. Fail. Anal.*, 2002, **9**, 99–107.
2. G. Kumar, S. Hegde and K. N. Prabhu: *J. Mater. Process. Technol.*, 2007, **182**, 152–156.
3. N. D. Alexopoulos and Sp. G. Pantelakis: *Mater. Des.*, 2004, **25**, 419–430.
4. S. B. Davenport, N. J. Silk, C. N. Sparks and C. M. Sellars: *Mater. Sci. Technol.*, 2000, **16**, 539–546.
5. P. S. Follansbee and U. F. Kocks: *Acta Metall.*, 1988, **36**, 81–93.
6. P. Wanjara, M. Jahazi, H. Monajati, S. Yue and J. P. Immarigeon: *Mater. Sci. Eng. A*, 2005, **A396**, 50–60.
7. S. Nemat-Nasser and J. B. Isaacs: *Acta Mater.*, 1997, **45**, 907–919.
8. Y. Niu, H. L. Hou, M. Q. Li and Z. Q. Li: *Mater. Sci. Eng. A*, 2008, **A492**, 24–28.
9. R. Kapoor, D. Pal and J. K. Chakravartty: *J. Mater. Process. Technol.*, 2005, **169**, 199–205.
10. M. Mulyadi, M. A. Rist, L. Edwards and J. W. Brooks: *J. Mater. Process. Technol.*, 2006, **177**, 311–314.
11. H. Sheikh and S. Serajzadeh: *J. Mater. Process. Technol.*, 2008, **196**, 115–119.
12. J. Luo, M. Q. Li and W. X. Yu: *Mater. Des.*, 2010, **31**, 3078–3083.
13. N. S. Reddy, Y. H. Lee, C. H. Park and C. S. Leea: *Mater. Sci. Eng. A*, 2008, **A492**, 276–282.
14. H. Kuhn and D. Medlin: in 'ASM handbook', Vol. 8, 'Mechanical testing and evaluation'; 2000, Materials Park, OH, ASM International.
15. 'Standard test methods for deglazing force of fenestration products', ASTM E987-88, Vol. 03.01, ASTM Philadelphia, PA, USA, 1988.
16. C. M. Estey, S. L. Cockcroft, D. M. Maijer and C. Hermesmann: *Mater. Sci. Eng. A*, 2004, **A383**, 245–251.
17. J. S. Ahmad and J. Twomey: *J. Mater. Process. Technol.*, 2007, **186**, 339–345.
18. L. H. Feng and J. Lu: *Expert Syst. Appl.*, 2010, **37**, 2974–2977.
19. K. M. Kanti and P. S. Rao: *J. Mater. Process. Technol.*, 2008, **200**, 300–305.
20. L. M. Almeida and T. B. Ludermi: *Neurocomputing*, 2010, **73**, 1438–1450.