Artificial neural network models for production of nano-grained structure in AISI 304L stainless steel by predicting thermo-mechanical parameters

F. Forouzan[1*], A. Najafizadeh[2], A. Kermanpur[3] and A. Hedayati[4]
Department of Materials Engineering,IsfahanUniversity of Technology,Isfahan 84156-83111,Iran

Abstract
An artificial neural network (ANN) model is developed for the analysis, simulation, and prediction of the austenite reversion in the thermo-mechanical treatment of 304L austenitic stainless steel. The results of the ANN model are in good agreement with the experimental data. The model is used to predict an appropriate annealing condition for austenite reversion through the martensite to austenite transformation. This model can also be used as a guide for further grain refining and to improve mechanical properties of the AISI 304L stainless steel.

Keywords: Metals and alloys, Phase transitions, Computer simulations.

1. Introduction
Austenitic stainless steels (ASSs) have good corrosion resistance and ductility. However, the yield strength of these steels is relatively low in the annealed state.[1,2)] There are various strengthening mechanisms for ASSs such as grain refining, transformation strengthening and work hardening while they cannot be hardened by heat treatment.[3,4)] However, grain refinement is the only method which improves both strength and toughness simultaneously.
Austenite phase in ASSs is not a stable phase in the general case and this fact holds particularly for the 300 steel series. Two types of martensite; namely, α′ (bcc, ferromagnetic) and ε (hcp, non-ferromagnetic), can be produced by cold deformation. The deformation and the content of strain-induced martensite depends upon the austenite stability (chemical composition and initial austenite grain size) and cold deformation conditions (the amount of deformation, deformation temperature, strain and strain rate).[4-9)]
One of the best methods to obtain ultra-fine austenite grains in ASSs is a thermo-mechanical process consisting of conventional cold rolling and annealing. In this process, the austenite phase should be almost completely transformed to strain-induced α′-martensite by heavy cold rolling at low temperatures, and subsequently strain–induced α′-martensite phase should be completely reverted to ultra fine austenite grains by annealing.
Key points in this thermo-mechanical process are as follows:Firstly, the strain-induced α′-martensite should be further heavily
deformed during cold rolling so that any lath α′-martensite structure is destroyed prior to reversion treatment.
This means that additional cold working after α′-martensite saturation leads to a higher capability for obtaining finer austenite grains.
Secondly, the strain-induced and deformed α′-martensite should be reverted to austenite at temperatures as low as possible to suppress grain growth of reverted austenite Fig.1.[3,10)]
Therefore, mechanical properties of austenitic stainless steel is a complex function of deformation temperature and annealing conditions, and it is very difficult to develop a complete mathematical model to predict the final properties of this steel after austenite reversion. An engineering approach to predict the properties of ASSs is based on utilization of artificial neural networks (ANNs). A neural network acts like a learning machine that enables us to model the hidden input– output relationships accurately. Several attempts have been made to predict the martensite content of these steels after cold work (by tensile test) and reversion of strain-induced α′-martensite to austenite using neural network.[11-13)]
In this work, the effects of annealing conditions on the reversion of α′-martensite to austenite in the cold rolled AISI 304L stainless steel at two cold rolling temperatures in a wide range of conditions are investigated and modeled as a guide map by ANNs.

* *Corresponding author:*
Tel: +98-311-3912752 Fax: +98-311-3915738
E-mail: f.forouzan@ma.iut.ac.ir
Address: Department of Materials Engineering,IsfahanUniversity of Technology,Isfahan 84156-83111,Iran
1,4. M.Sc.
2. Professor
3. Associate Professor

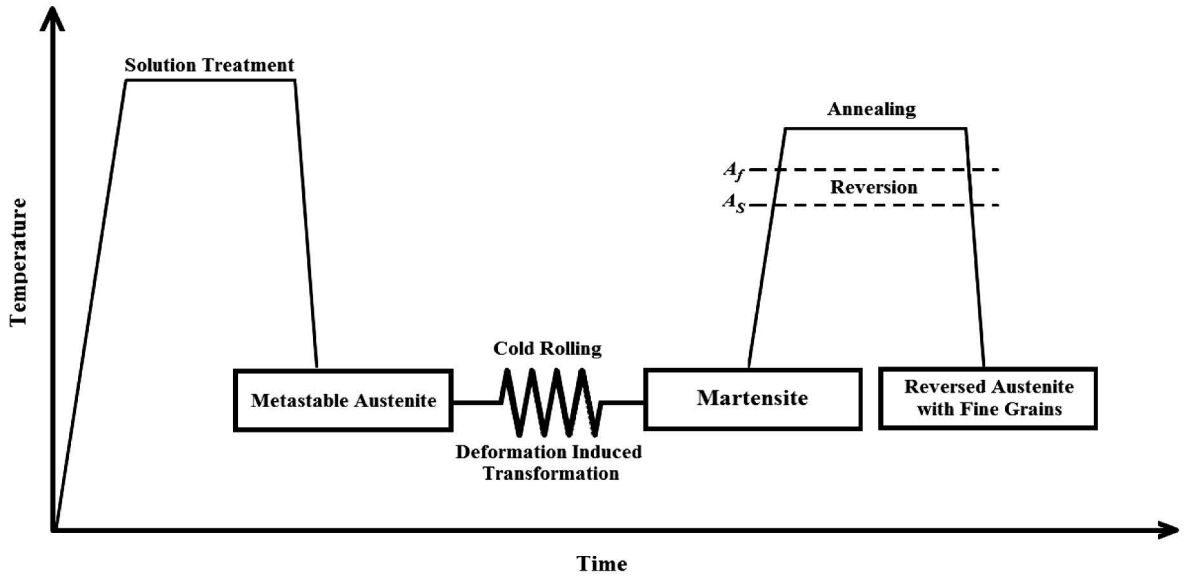

Fig.1. Thermo-mechanical treatment to achieve ultrafine grains in metastable austenitic stainless steels. A_s and A_f are the reversion start and finish temperatures, respectively.

2. Materials and Methods

The chemical composition of the studied industrial steel sheet is given in Table 1. The sheet material was cold rolled at -15 and 0°C with inter-pass cooling up to 90%. Deformed specimens were then annealed at temperature ranges of 650–900 °C from 5 seconds to 8 hours.

Specimens for α′-martensite measurements were cut from the strained sheets with an abrasive cutting machine and the edges of the specimens were finished with SiC emery paper. A Ferritscope (Helmut Fischer GmbH, model MP30E) was used for the quantification of α′-martensite phase.

The device was calibrated with δ-ferrite standard specimens and the results were converted to the α′-martensite contents.[14] The microstructures were studied using a Scanning Electron Microscope (SEM Philips X230).Prior to this study, the specimen surfaces were electropolished by using an electrolyte (200 ml perchloric acid, 800 ml ethanol) at 30 V for 30 s. The specimens were electroetched with an electrolyte at 20 V for 30 s. Electrolytic etching with a mixture of 60 ml nitric acid and 40 ml distilled water was used to reveal the austenite grain boundaries. The etching was carried out at 1.0 V for about 8 minutes.

3. Theory/Calculation

ANN modeling follows these steps: determination of input/output parameters; collection of data; analysis of the data; training and testing of the neural network; and using the trained network for simulation and prediction. Model training includes the choice of architecture, training algorithms and parameters of the network.[15]

3.1. Database

In the case of cold-rolled steel sheets, the most important parameters of heat treatment, which dictate microstructure and mechanical properties of ASSs, are the annealing time and temperature used as the input data.

A total number of 140 data based on thermo-mechanical tests were considered for the analysis. Cold reduction was 90% and the rolling temperature was -15 and 0 °C (all of them below M_d temperature).These values were called subzero and zero temperatures, respectively. Data were divided into three groups, 50% for the training set, 25% for the validation set, and 25% for the test set. In this study, the validation set was used to control the training process and the testing set was used to evaluate the generalization ability of the trained network. The volume fraction of strain-induced martensite values used as output data in two networks were normalized to bound the values to the set [0, 1] using the following formula:

$$X_N = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Where X_N is the normalized value of variable X with maximum and minimum values given by X_{max} and X_{min}.

Table 1. Chemical composition of 304L austenitic stainless steel (Weight Percent).

Type	C	Si	Mn	Cr	Ni	Cu	Mo	Nb	Fe
304L	0.0269	0.427	1.58	18.2	8.22	0.58	0.348	0.0020	base

3.2. Network architecture

Feed forward networks often have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. For multiple-layer networks the number of layers determines the superscript on the weight matrices.[16)]

Figure 2 represents the structures of neural networks modeled in this work schematically. As Hornik et al.[17)] have shown, an ANN with one hidden layer with sigmoid transfer function (logarithm sigmoid or hyperbolic tangent)can approximate any function, the three-layered neural network (one input, one hidden, and one output layer) was used in the present work for both of the models. Often the back-propagation method is used to train ANNs, in which the gradient is computed for nonlinear multilayer network.

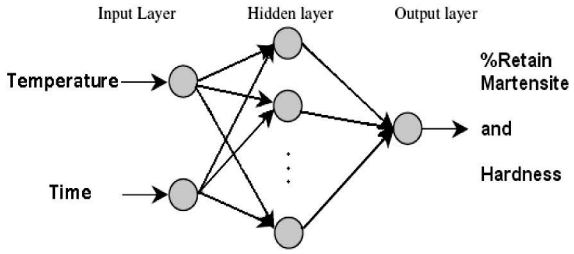

Fig. 2. Neural network architectures modeled in this work.

3.3. Network training

The neural networks used in our study are the supervised multilayered feed-forward back-propagation network trained with Levenberg – Marquardt (LM) algorithm which is the fastest training algorithm for networks of moderate size.[18)]

The inputs p_i are multiplied by weights w_{ji} for a hidden node n_j. Then, the weighted sum of $w_{ji}p_i$ is added to a bias value b_{ji} and finally operated by a suitable transfer function (f). The operation can be written as:

$$n_j = f\left(\sum w_{ji} p_i + b_{ji}\right) \qquad (2)$$

A hyperbolic tangent function was used here as the transfer function as follows:

$$f(n) = \frac{1-\exp(-\alpha p_i + c)}{1+\exp(-\alpha p_i + c)} \qquad (3)$$

Where α and c are constants. In order to find out the appropriate network architecture,similar operations are repeated for the hidden layers. The hidden layers contribute to the neuron output through a linear function. The output can be written as:

$$a = \left(\sum w_j n_j + b'\right) \qquad (4)$$

Where w_j and b′ are new sets of weight and bias values. The appropriate notation is used in the two-layered tansig/purelin network shown in Fig 3.

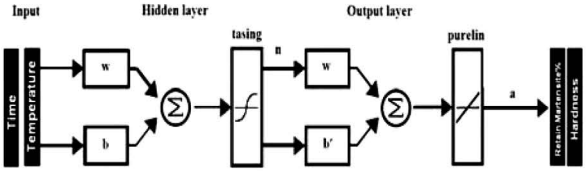

Fig. 3. Schematic diagram of the ANN model for prediction of properties of AISI304L steel after annealing.

4. Result and discussion

4.1. Model performance

It is usefulto plot the training, validation, and test errors in order to check the progress of training. To make sure that the network does not end at a local minimum value, it is necessary to plot the error function versus the number of epochs. This is shown in Fig. 4 where the error is represented by the mean square errors (MSE) which is presented as:

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

Where t_i is the target output and y_i is the network output.

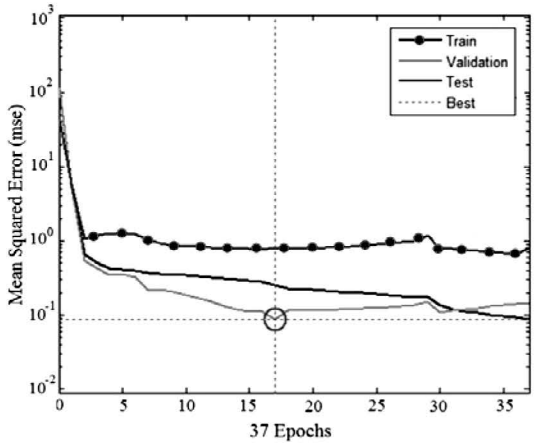

Fig. 4. Error function versus the number of epochs to schematic representation of the stopping technique used in this work.

As the number of cycles increases, the error decreases until about 17 cycles and then it increases. Different training options, such as training on variations of mean square error for better generalization, training against a validation set, and training until the gradient of the error reaches a minimum, have been attempted in order to get the best result.

Superior performance of the ANN model was achieved by changing the parameters of modeling described above. The performance of the resulting

model on different datasets, namely, training, validation, testing, and the entire data are shown in Fig. 5. It shows the linear regression between the network output and the corresponding data for subzero reversion network model. For all other trained networks, the outputs trace the targets very well. The main source of deviation was the mentioned contradiction in the experimental data.[19)]

Table 2 shows that similar performance is achieved for another network used for the simulation of the reversion of zero specimens. As can be seen,R-values(The R-value relates the goodness of the fit of the line. R-values closer to +1 or -1 indicate a better fit than values closer to 0.)for all cases of training, validation, and test data sets were above 0.95.

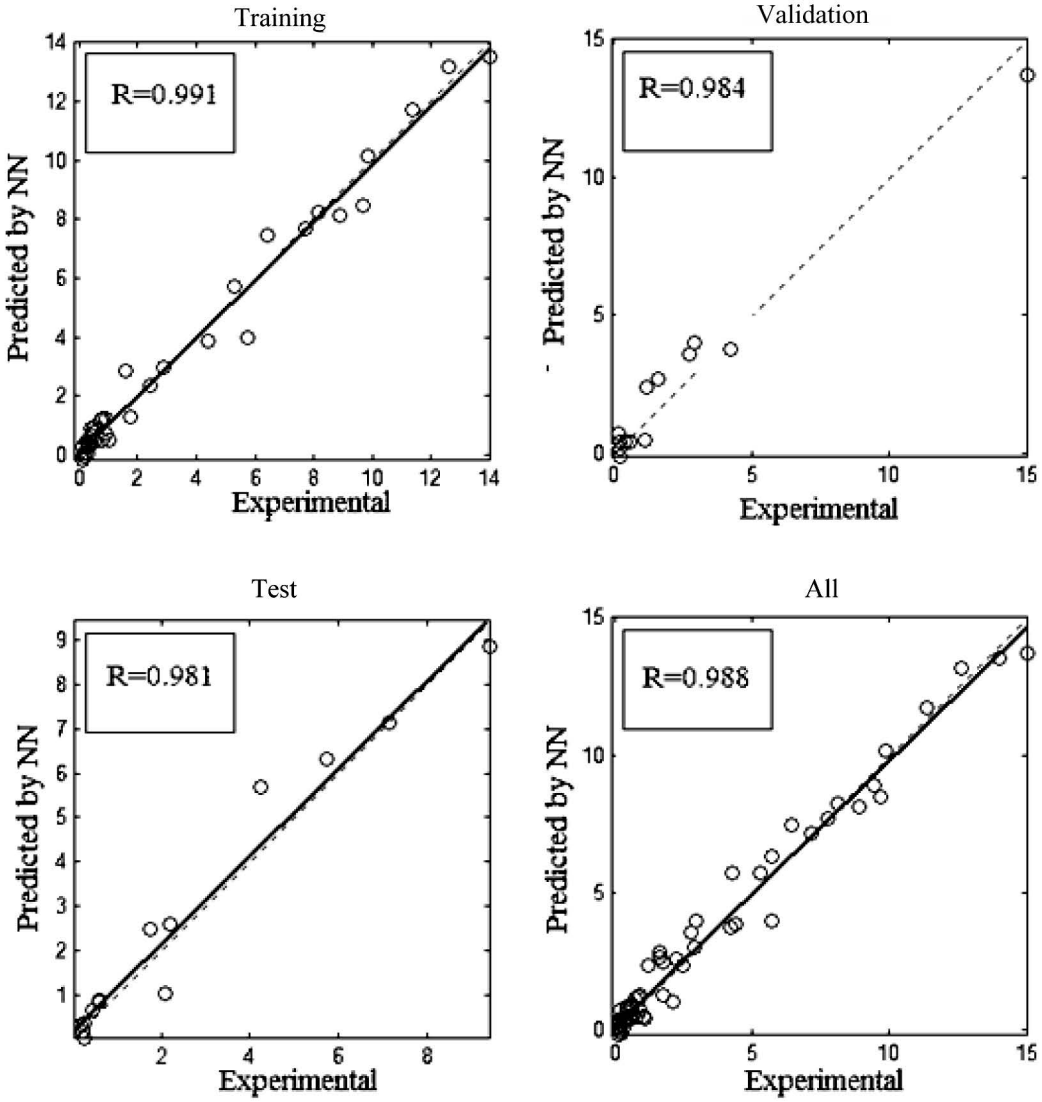

Fig. 5. Performance of ANN Subzero Reversion model for training, validation, testing data and all data, respectively.

Table 2. Network results for training algorithms used in this work.

Neural Network	The best architecture	Mean square error	R in linear regression
Subzero reversion model	2-5-1	0.25	0.98
Zero reversion model	2-6-1	0.307	0.97

4.2. Modeling of reversion

Reversion of austenite from α′-martensite has been analyzed during annealing at different temperatures and times. To increase the mechanical properties of this type of steel, its microstructure after annealing must be fully austenitic with nano/submicron grains sizes. Thus, the temperature and time of annealing must be designed as low and short as possible to prevent grain growth of austenite after reversion treatment. Figure 6(a) shows the ultra fine austenitic structure after adequate annealing, and Figure 6(b) shows the grain growth in the structure after extra annealing.

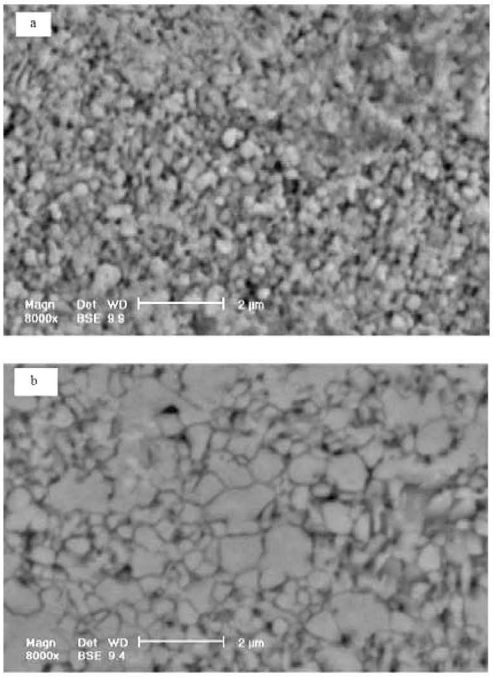

Fig. 6. SEM image of the AISI 304L stainless steel specimen after 90% cold reduction at 0˚C.(a) annealed at 800 °C for 2min (b) annealed at 800 °C for 14min.

The effect of annealing conditions on the percentage of austenite reversion for both rolling temperatures by using the results of the developed networks is shown in Fig. 7.a and b. These figures show that increasing the annealing temperature and duration lead to increased martensite reversion. Furthermore, reversion rate increases with the annealing temperature. As can be seen in these figures, at high annealing temperatures, the reversion rate at short annealing time on the reversion process will be which reversion is complete but it is not too long to allow grain growth.periods of annealing is high but rapidly descends by increasing the annealing time so that the effect of minor, but it is important for growing the austenite grain size. Therefore, the optimal point is the time at which reversion is complete but it is not too long to allow grain growth.

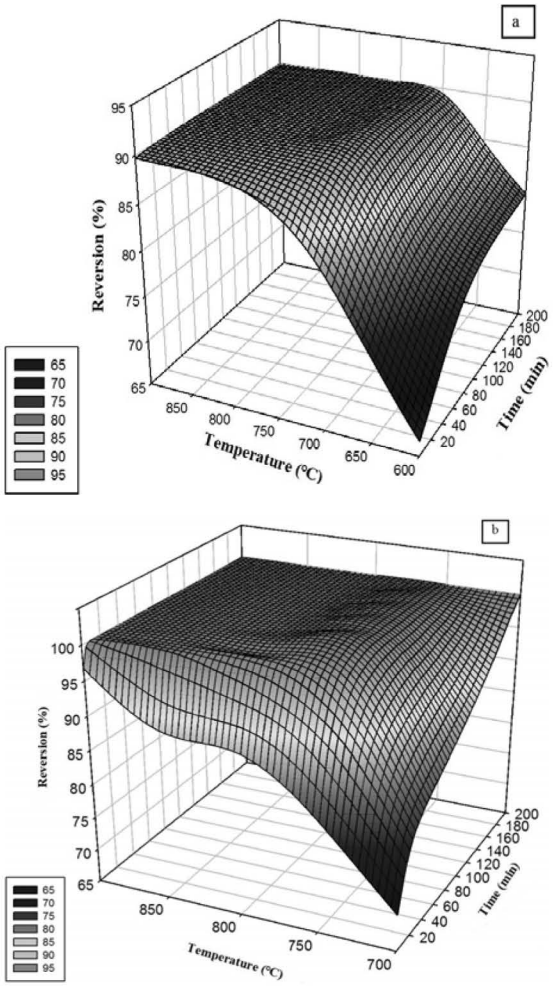

Fig. 7. The effect of annealing temperature and time on the relative percentage of reversion (a) cold rolled at -15ºC, (b) cold rolled at 0 ºC.

Contour lines of the percentage of reversed austenite after annealing are illustrated in Fig. 8. These figures show that the complete reversion at any acceptable time is only achievable by annealing at temperatures higher than 750 ˚C. As can be seen, in lower temperatures,e.g. 600 and 650˚C, the time to reach up to 90% austenite is very long.

Comparison of rolling at subzero and zero reversion curves as a function of annealing conditions reveals the following results:

- In zero temperature rolled specimen, percentage of the reverted austenite was not adequate in a short time (until 20 minutes) although the temperature was very high, but in subzero temperature rolled specimen at temperatures above 800˚C, reversion was approximately completed in a very short time (below 10 seconds).
- The high gradient of 90%-contour-line in zero temperature rolled diagram implies that increasing the annealing temperature up to 750 ˚C does not decrease the annealing time.

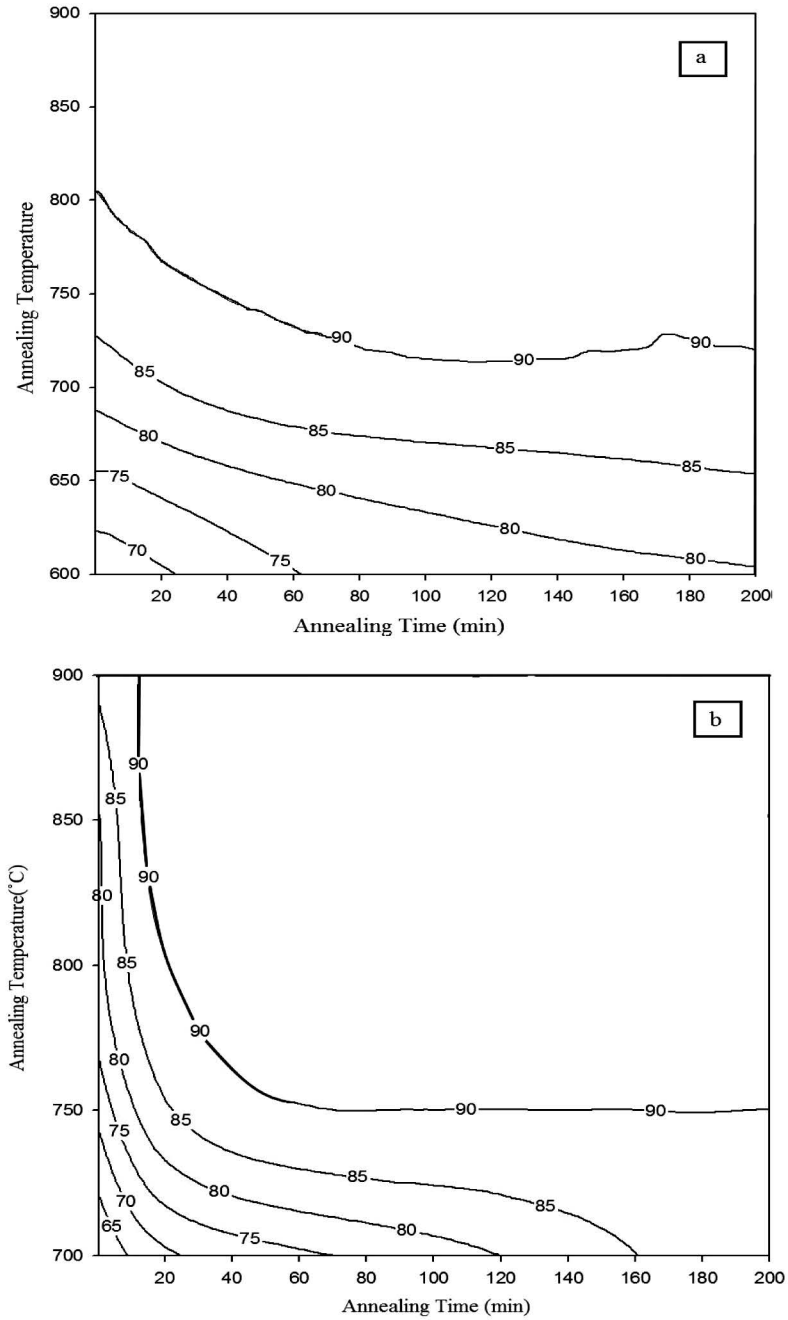

Fig. 8. Contour lines of relative percentage of reversion as a function of annealing temperature and time (a) coled rolled at -15ºC, (b) coled rolling at 0 ºC.

However, above 750 ˚C temperatures, the reversion was almost complete after 30 minutes. The experimental retained martensite data measured by Ferritescopeare also presented by points in Fig 9. As expected, these two data series agree with each other.

The results presented here showed that the ANN models can predict the influences of the annealing time and temperature on retained martensite in the structure. The ANN models can also be used as a guide for prediction of optimum annealing time and temperature according to the required conditions.

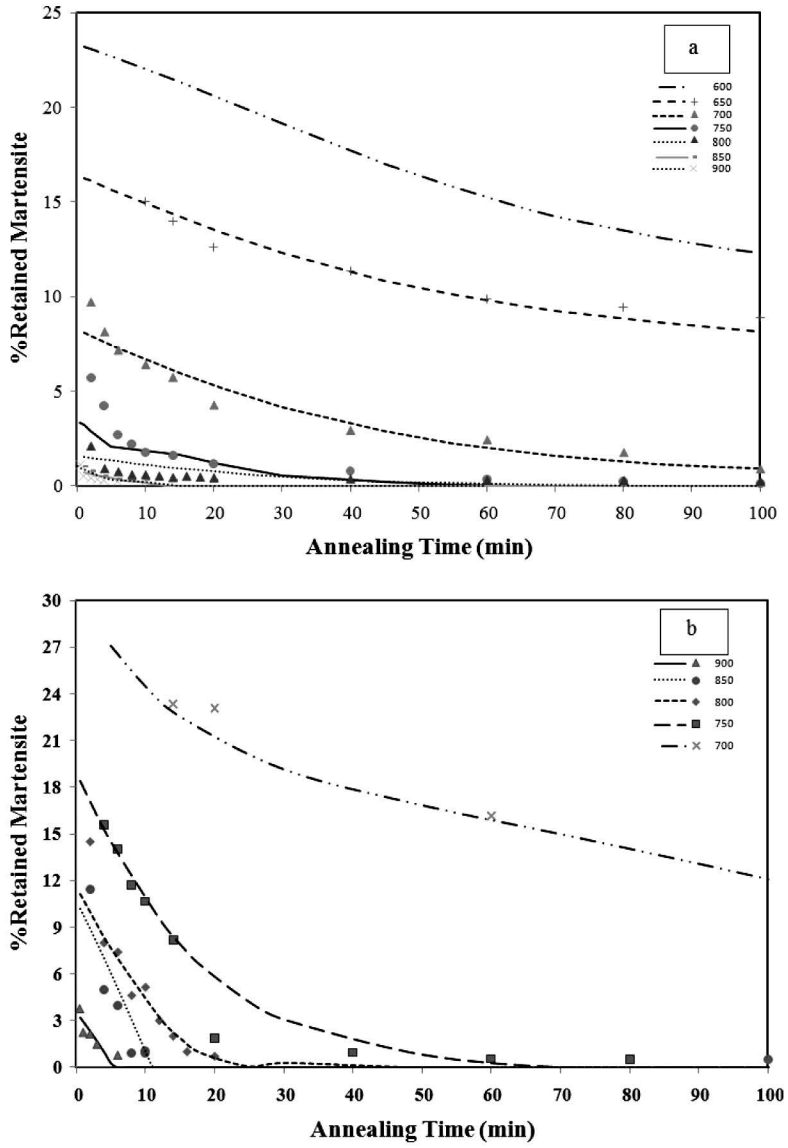

Fig. 9. Comparison of ANN model calculations with experimental values of annealing conditions after (a) cold rolling at -15˚C (b) cold rolling at 0˚C.

5. Conclusions

The effect of rolling temperature and annealing condition on the reversion of strain-induced martensite to austenite was modeled for AISI 304L stainless steel by means of artificial neural networks. The developed model showed that higher annealing temperature and time resulted in a greater reversion martensite to austenite. Furthermore, the complete reversion at any acceptable time was only achievable by annealing at temperatures higher than 750°C. Finally, increasing the rolling temperature leads to a much longer annealing time for complete reversion.

References

[1] The Materials Information Society: ASM Specialty Handbook: Stainless Steel, ASM INTERNATIONAL, Metals Park, OH, (1994), 314-315.

[2] A.F. Padilha, R.L. plaut and P.R. Rios: ISIJ Int., 43(2003), 135.

[3] T. Maki: Stainless steel: Curr. Opin. Solid State Mater. Sci., 2(1997), 290.

[4] Y. Murata, S. Ohashi and Y. Uematsu: ISIJ Int., 33(1993), 711.

[5] A.F. Padilha and P.R. Rios: ISIJ Int., 42(2002),325.

[6] G.B. Olsen and M. Cohen:J. Less-Common Met., 28(1972), 107.

[7] G.B. Olsenand M. Cohen: Metall. Trans. A, 6(1975), 791.

[8] P.L. Mangonon and G. Thomas: Metall. Trans, 1(1970), 1577.

[9] P.L. Mangonon and G. Thomas: Metall. Trans. A, 1(1970), 1587.

[10] K.Tomimura, S. Takaki, S. Tanimoto and Y. Tokunaga: ISIJ Int., 31(1991), 721.

[11] H.Mirzadeh and A.Najafizadeh: Mater. Charact., 59(2008), 1650.
[12] H. Mirzadeh and A. Najafizadeh: J. Alloys Compd., 476(2009), 352.
[13] H. Mirzadeh and A. Najafizadeh: Mater. Des., 30(2009), 570.
[14] J. Talonen, P. Aspegren and H. Hänninen: Mater. Sci. Technol., 20(2004),1506.
[15] Z. Guo and W. Sha: Comput. Mater.Sci., 29(2004), 12.
[16]http://www.mathworks.com/access/helpdesk/help/pdf_doc/nnet/nnet.pdf/
[17] K. Hornik, M. Stinchcombe and H. White: Neural Network, 2(1989), 359.
[18] C.M. Bishop: Neural Networks for Pattern Recognition, Clarendon Press, Oxford,(1995).
[19] S. Malinov, W. Sha and Z. Guo: Mater. Sci. Eng. A, 283(2000),1.