



J. Serb. Chem. Soc. 77 (7) 937–944 (2012)
JSCS–4322

Application of a Bayesian Artificial Neural Network and the Reversible Jump Markov Chain Monte Carlo Method to predict the grain size of hot strip low carbon steels

MOHSEN BOTLANI-ESFAHANI^{1*} and MOHAMMAD REZA TOROGHINEJAD²

¹Department of Materials Engineering, Lenjan Branch, Islamic Azad University, Isfahan, Iran and ²Department of Materials Engineering, Isfahan University of Technology, Isfahan, 84156-83111, Iran

(Received 11 November 2011, revised 31 January 2012)

Abstract: An Artificial Neural Network (ANN) with Reversible Jump Markov Chain Monte Carlo (RJCMC) simulation was used to predict the grain size of hot strip low carbon steels, as a function of steel composition. The results show good agreement with experimental data taken from the Mobarakeh Steel Company (MSC). The developed model is capable of recognizing the role and importance of elements in grain refinement. Furthermore, the effects of these elements, including manganese, silicon and vanadium, were investigated in the present study, which were in good agreement with the literature.

Keywords: artificial neural network; grain size; hot strip; low carbon steel.

INTRODUCTION

Grain size is an important aspect of microstructure with respect to mechanical properties of steels. The ferrite in low carbon steels is typically strengthened by grain refinement, precipitation hardening, and, to a lesser extent, solid-solution strengthening. Grain refinement is the most desirable strengthening mechanism because it improves not only strength but also toughness.¹ According to Equation (1), indicating the Hall–Petch Relation, fine grain sizes produce higher yield strengths, (σ_{yield}):²

$$\sigma_{\text{yield}} = \sigma_{\text{init}} + k_y d_{\alpha}^{-1/2} \quad (1)$$

where σ_{init} is the yield strength for a polycrystalline material, k_y is a constant and d_{α} is a measure of the ferrite grain size. Grain size also has an effect on the ultimate tensile strength by changing the work-hardening rate. Work-hardening occurs within the grains during plastic deformation according to Morrison:³

* Corresponding author. E-mail: botlani@iauln.ac.ir
doi: 10.2298/JSC111115011B

$$n = 5 / (10 + d_{\alpha}^{-1/2}) \quad (2)$$

where n is work-hardening exponent and d_{α} is the grain size.

The main purpose of hot stripping is to refine the microstructure of steels, thereby enhancing both strength and toughness.⁴ Ferrite mainly nucleates at the austenite grain boundaries and thus finer austenite grains results in finer ferrite grains. Further ferrite refinement could be achieved by transformation from deformed austenite grains because deformation increases the ferrite nucleation rate.²

Precise prediction of grain size and properties of hot stripped steel products depends on the thermo-mechanical behavior of the steel, microstructure evolution and phase transformation during hot rolling stages and the cooling period. These are complicated metallurgical phenomena and strongly depend on chemical composition, therefore development of a physical model to analysis these parameters and predict the final grain size is cumbersome. In addition, the accuracies of the hitherto developed models are somehow questionable and not suitable for practical purposes. On the other hand, estimating the final grain size of low carbon steel strips in terms of chemical composition is desirable from an engineering viewpoint. In the present work, by selecting more relevant inputs and using a hybrid Bayesian Artificial Neural Network (ANN) Model assisted with the Reversible Jump Markov Chain Monte Carlo (RJCMC) Method, the final grain size in low carbon steel strips was predicted.

METHODS

Artificial Neural Network (ANN)

An artificial neural network model maps sets of input data onto a set of appropriate output data. It is a modification of the standard linear perceptron in that, it uses three or more layers of neurons with nonlinear activation functions. This makes it more powerful than the perceptron, and can distinguish data that is not linearly separable.⁵ The basic ANN Model with k outputs is:

$$f_k(X, W) = w_{k0} + \sum_{j=1}^m w_{kj} \tanh(w_{j0} + \sum_{i=1}^d w_{ji} x_i) \quad (3)$$

where X is a d -dimensional input vector, W denotes the weights, and the indices i and j correspond to the input and hidden units, respectively.⁶ The schematic architecture of an ANN Model is sketched in Fig. 1.

Traditionally, the complexity of an ANN is controlled with early stopping, whereby part of the data is used to train the network and the other part is used to control the complexity of the model.

Bayesian learning for an ANN

Early stopping is inefficient because the effective complexity may be much less than the number of parameters in the model. Within the Bayesian framework, the weights of the net-

work are considered as random variables and the posterior distribution of the weights updated according to the Bayes' Rule:⁷

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}} \quad (4)$$

This equation in terms of Artificial Neural Networks is:

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)} \propto L(\theta|D)p(\theta) \quad (5)$$

where $p(\theta)$ is the prior distribution for the model parameters θ , $D = \{(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), \dots, (\mathbf{x}^{(n)}, \mathbf{y}^{(n)})\}$ is the observed data and $L(\theta|D)$ is the likelihood function that gives the probability of the observed data as a function of the unknown model parameters.⁶

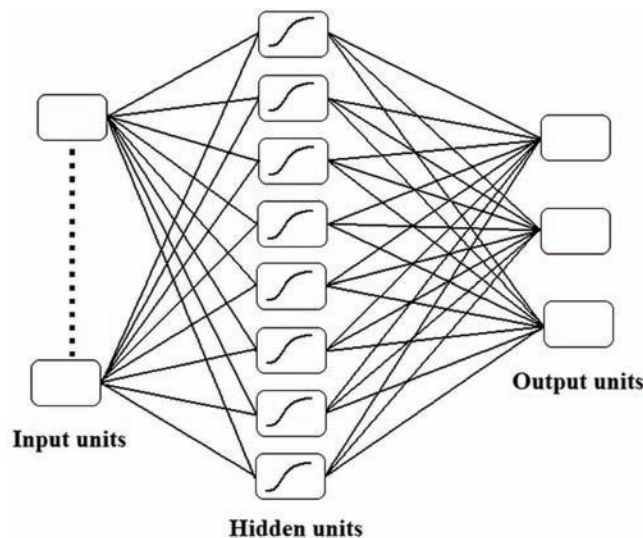


Fig. 1. Schematic structure of the ANN Model.

Reversible Jump Markov Chain Monte Carlo Method

Neal introduced an implementation of Bayesian learning for an ANN in which difficult integrations accompanied within this framework were performed using the Markov Chain Monte Carlo (MCMC) Method. In this application, samples (in model parameters space) are generated using the Markov Chain Monte Carlo Method to estimate the desired posterior distributions.⁶

In practical problems, as in the present study, it is usual to measure many variables, but it is not necessarily known which one of them is relevant and is required to solve the problem. To make the model more explainable or to reduce the measurement cost and computation time, it may be useful to select a model with a smaller set of input variables.⁶ Consequently, the RJMCMC method is applied for this modeling. This algorithm allows jumps between models with different dimensional parameter spaces with respect to the number of inputs chosen in the model. The RJMCMC visits the models according to their posterior probability, which allows it to be used for model selection.⁸

Modeling database

Since an ANN Model is empirical, its performance depends on the dataset used for training. In this research, the data was taken from the Mobarakeh Steel Company (MSC) and consisted of 624 metallographic images. At this company, these images were classified into three groups according to ASTM (E-112), a standard that assigns larger numbers to finer grain structures. An example of such a database is shown in Fig. 2. Further information is also given in Table. I. The input parameters are chemical composition of the strips, which includes 14 elements. Additional input variables are given in Table. II.

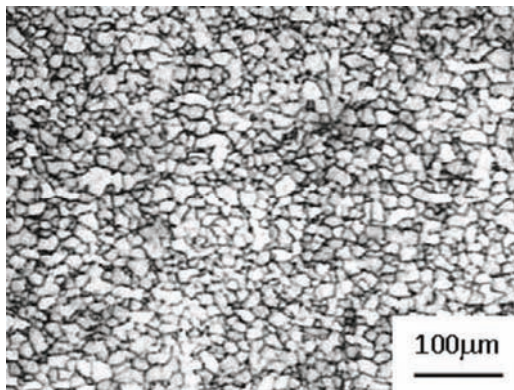


Fig. 2. One sample of data, microstructure of a produced steel with ASTM grain size No. 9.0.

TABLE I. Output data information

Number of samples	ASTM (E-112) grain No.
162	8.5
294	9
167	9.5

Model training

The modeling database was divided into training and test sets, which included 60 % and 40 % of the data, respectively. Training was started with a (14–8–3) architecture and model selection procedure was evaluated by an internal procedure of the RJMCMC algorithm, as mentioned above. The result of this training indicated a chain of network parameters. When this chain converged into a stable distribution, a sample of the chain (network parameters) was selected based on the minimum classification error of the model on the test dataset. This modeling was implemented by MCMC methods for Multilayer Perceptron (MLP) Network and Gaussian Process (GP).*

RESULTS AND DISCUSSION

Performance of the model

Calculation of the misclassification error on test data is a popular way to show the prediction accuracy (generalization) of a classifier model. This error is calculated according to:

* www.lce.hut.fi/research/mm/mcmcstuff/

$$\text{Misclassification error} = \frac{\sum |\text{Test data} - \text{Model result}|}{\text{Number of test data}} \times 100 \quad (6)$$

The achieved model revealed a misclassification error of just 2.439 %, which is very low and indicates that the model has good generalization. More information about the misclassified error is available in Table. III.

TABLE II. Input parameter information

No.	Inputs	Min.	Max.	Mean	SD
1	C / wt. %	0.032	0.179	0.1272	0.0312
2	Si / wt. %	0.008	0.218	0.0637	0.0798
3	Mn / wt. %	0.191	1.15	0.6466	0.211
4	P / wt. %	0.002	0.025	0.0072	0.0022
5	S / wt. %	0.001	0.02	0.0089	0.0029
6	Cu / wt. %	0.004	0.078	0.03	0.0109
7	Al / wt. %	0.015	0.075	0.0454	0.0119
8	N / ppm	16	75	38	8.8
9	Nb / wt. %	0	0.045	0.0051	0.0105
10	V / wt. %	0	0.011	0.003	0.0014
11	Ti / wt. %	0	0.042	0.0017	0.0031
12	Mo / wt. %	0	0.019	0.0038	0.0045
13	Cr / wt. %	0.004	0.194	0.0131	0.012
14	Ni / wt. %	0.02	0.042	0.03	0.0034

TABLE III. Misclassified case for the number of misclassified 2

Test target data	Model result
9	8.5
9.5	9

Model results

As mentioned, the RJMCMC method can select potentially useful inputs according to the marginal probabilities of the inputs. The result of this analysis indicated the importance of the contents of Si, Mn and C on grain refinement, which were significantly greater than the concentrations of other elements. The most effective element for grain refinement was recognized to be that of vanadium. However, its concentration in these steels was very low.

For testing, the results of the model are depicted when the concentrations of elements are on their mean values, which are given in Table II and the micro-alloying elements (*i.e.*, Nb, Ti and V) are not present. The model result of this analysis is shown in Fig. 3. Manganese stabilizes austenite, therefore decreases the austenite to ferrite transformation temperature and hence refines the grain structure. In addition, manganese can enhance the precipitation strengthening of the vanadium micro-alloyed steels and to a lesser extent, niobium micro-alloyed steels.¹

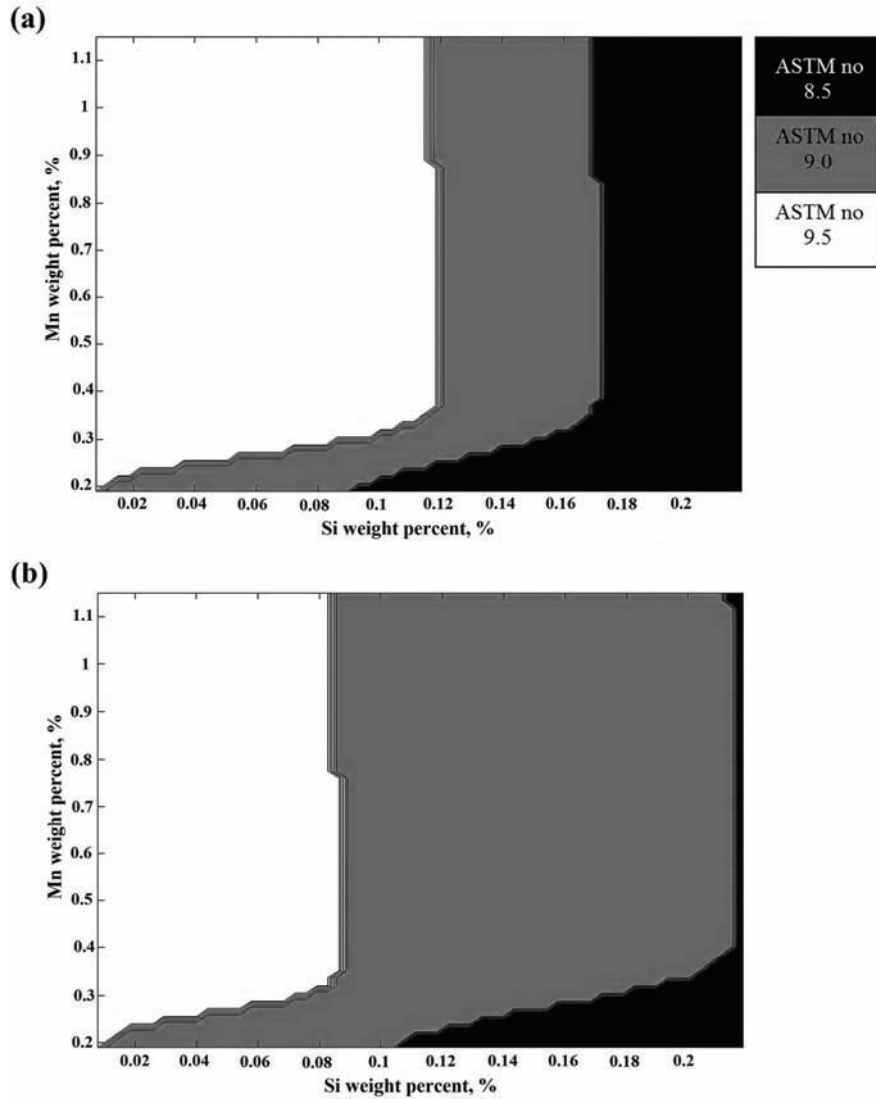


Fig. 3. Model results in respect to the silicon and manganese concentration in 0.015 wt. % C and 0.035 wt. % Al. a) Absence of micro-alloying elements (Nb, V and Ti). b) Minor addition of vanadium (0.008 wt. %) in the absence of the other micro-alloying elements (Nb and Ti).

Fig. 3a reveals the determining role of silicon on grain size in the absence of micro-alloying elements (*i.e.*, Nb, Ti and V). The figure shows that the silicon concentration divides the figure into three regions, including finer, mild and coarser grain structures. This figure also indicates that increasing Si content increases the grain size. This is because silicon is a ferrite stabilizer and promotes ferrite grain growth.⁹ Fig. 3b shows that addition of small amount of vanadium,

0.008 wt. %, to steel severely contracts the coarser grain region. Vanadium acts as a scavenger for oxides, and forms nano-scale inter-phase precipitations. This is mainly due to the rapid rate of the austenite to ferrite transformation, which produces these nano-scale precipitates.¹⁰ Furthermore, the addition of vanadium also reduces the finer grain area somewhat. This is because vanadium is strong carbide former and as the majority of such elements are ferrite stabilizers, it therefore promotes ferrite grain growth.¹¹ The net effect of this minor vanadium addition is to decrease the sensitivity of the grain size to the silicon content, and also reduction of coarse grain area.

CONCLUSIONS

A Bayesian ANN Model assisted by RJMCMC is capable of predicting the grain size of hot strip low carbon steels and can be used as a function of steel composition. The results are shown to be consistent with experimental data (acquired from Mobarakeh Steel Company data). Furthermore, the model recognizes the effects of relevant elements in grain refining. These are manganese, silicon and vanadium. The silicon concentration has a determining role, this effect has not been reported in the literature, and vanadium has a great impact on grain refining phenomena. The major advantage of this method, which was validated by literature in the present research, is the selection of useful inputs in complex problems with many inputs. As many problems in materials science and engineering are similar, this method could be useful for solving them.

Acknowledgment. . The authors wish to thank, Lenjan Branch, Islamic Azad University, and the Iranian Nanotechnology Initiative Council (INIC) for their supports as well as Isfahan Mobarakeh Steel Company for supplying experimental data.

ИЗВОД

ПРИМЕНА БАЈЕСОВИХ ВЕШТАЧКИХ НЕУРОНСКИХ МРЕЖА (ANN) И РЕВЕРЗИБИЛНЕ МОНТЕ КАРЛО МЕТОДЕ НА БАЗИ МАРКОВЉЕВИХ ЛАНАЦА (RJMCMC) ЗА ПРЕДВИЂАЊЕ ВЕЛИЧИНА ЗРНА У ТОПЛО ВАЉАНИМ ТРАКАМА ОД НИСКО УГЉЕНИЧНОГ ЧЕЛИКА

MOHSEN BOTLANI-ESFAHANI¹ и MOHAMMAD REZA TOROGHINEJAD²

¹Department of Materials Engineering, Lenjan Branch, Islamic Azad University, Isfahan, Iran u ²Department of Materials Engineering, Isfahan University of Technology, Isfahan, 84156-83111, Iran

Вештачке неуронске мреже (ANN) и реверзибилна Монте Карло Метода на бази Марковљевих ланаца (RJMCMC) су коришћени да би се предвидела величина зрна топло ваљаних трака од ниско-угљеничног челика у функцији њиховог састава. Резултати показују добро слагање са експерименталним подацима добијеним у *Mobarakeh* фабрици за производњу челика (MSC). Развијени модел има могућност идентификовања улоге и важности елемената за рафинацију зрна. Поред тога, ефекти ових елементата, укључујући манган, силицијум и ванадијум, су испитивани у овом раду и показују добру сагласност са подацима из литературе.

(Примљено 11. новембра 2011, ревидирано 31. јануара 2012)

REFERENCES

1. Keytosteel.com, *Control of high strength low alloy (HSLA) steel properties*, <http://www.keytosteel.com> (accessed in 2010)
2. S. V. Parker, *PhD Thesis*, University of Cambridge, UK, 1997
3. J. Ryu, *M. Sc. Thesis*, Pohang University of Science and Technology, Korea, 2008
4. I. Tamura, C. Ouchi, T. Tanaka, H. Sekine, *Thermomechanical processing of high strength low alloy steels*, Butterworth, London, UK, 1988
5. S. Haykin, *Neural networks: a comprehensive foundation*, 2nd ed., Prentice Hall International, Upper Saddle River, NJ, 1998
6. J. Lampinen, A. Vehtari, *Applied Computational Intelligence to Engineering and Business*, K. Wang, J Grundespenkis, A. Yerofeyev, Eds., Riga Technical University, Latvia, 2001, p. 7
7. M. Xu, G. Zeng, X. Xu, G. Huang, R. Jiang, W. Sun, *Water Air Soil Pollut.* **172** (2006) 167
8. A. Vehtari, J. Lampinen, *Neural Comput.* **14** (2002) 2439.
9. M. Umemoto, Z. G. Liu, K. Masuyama, K. Tsuchiya, *Sci. Mater.* **45** (2001) 391
10. H. K. D. H. Bhadeshia, R. W. K. Honeycombe, *Steels Microstructure and Properties*, 3rd ed., Elsevier, Oxford, UK, 2006
11. Y.-B. Zhang, D. Y. Ren, *Mater. Sci. Technol.* **19** (2003) 1029.