JMEPEG (2003) 12:623-628

Neural Network Modeling for the Prediction of Texture Evolution of Hot Deformed Aluminum Alloys

P. Barat and P.J. Withers

(Submitted 15 June 2003; in revised form 5 September 2003)

Commercial aluminum rolling mills operate under very restricted thermomechanical conditions determined from experience and plant trials. In this paper we report results for four-stand tandem mill rolling simulations within and beyond the thermomechanical conditions typical of a rolling mill by plane strain compression (PSC) testing to assess the effect of deformed conditions on the texture of the hot deformed aluminum strip after annealing. A neural network modeling study was then initiated to find a predictive relationship between the observed texture and the thermomechanical parameters of strain, strain rate, and temperature. The model suggested that temperature is the prime variable that influences texture. Such models can be used to evaluate optimal strategies for the control of process parameters of a four-stand tandem mill.

Keywords aluminum alloy, Gaussian process model, neural network, plane strain compression, rolling

1. Introduction

The majority of semifabricated, flat rolled aluminum products are manufactured using large and expensive hot and cold rolling facilities to meet the high standards required for subsequent downstream processing and in-service performance. For many products the hot rolling stage is crucial in determining the correct microstructure, including crystallographic texture, prior to cold rolling and in some cases, after final annealing. On-line monitoring for the control of metallurgical parameters is in its infancy within the industry, and strict practices for process control are enforced to achieve the correct metallurgy. These practices are traditionally established on the basis of experience and plant trials. To establish new processes it is always helpful to have physical simulations of the process as an alternative to costly plant trials. As a result, in recent years the determination of process-microstructure relationships relating to hot tandem mill rolling of aluminum alloys has relied on extensive use of laboratory plane strain compression (PSC) testing.[1]

The microstructural changes, especially with respect to crystallographic texture, that occur during hot tandem mill rolling of aluminum and aluminum alloys are very sensitive to strain, strain rate, and temperature. These changes include both dynamic (during rolling) and static (between rolling passes and after rolling) structural changes and involve recovery, recrystallization, and grain growth. Although many experimental studies have been conducted to examine the microstructural evolution during hot deformation and subsequent recrystallization for aluminum alloys, the quantitative link between microstructural changes and thermomechanical history has been studied in a very limited manner and only for specific alloys.[2-5]

P. Barat (presently at Variable Energy Cyclotron Centre 1/AF, Bidhan Nagar Kolkata-700 064, India) and **P.J. Withers**, Materials Science Centre University of Manchester and UMIST, Grosvenor St., Manchester, M17 HS, UK. Contact e-mail: pbarat@veccal.ernet.in.

Control of texture during the manufacture of aluminum strips is particularly important for the sheet used to make beverage can bodies. The alloy used for this process is AA3104, the composition of which, based on Al-Mn-Mg with impurity levels of Fe and Si and deliberate additions of Cu, is tailored to provide the correct balance of strength, formability, and ability to be recycled. Texture control is necessary for this material during rolling to minimize the formation of "ears" in the drawn cups during forming of the can body from the final gauge sheet.[1]

The tandem mill operating conditions impose a particular thermomechanical history of strain, strain rate, and temperature on the strip as it passes through the mill. Commercial rolling mills are operated within a controlled range of process conditions determined from experience and from plant trials. In this work, a set of four-stand tandem mill rolling simulations have been performed both within and beyond the extremes of thermomechanical history that might be experienced during commercial processing to assess the effect on the texture of the hot deformed strip after annealing. A modeling study was then initiated to find a predictive relationship between the observed texture and the thermomechanical parameters of strain, strain rate, and temperature.

For many years there have been attempts to predict microstructural evolution in thermomechanically processed metals based on basic physical metallurgical principles.[6-9] However, while progress has been significant, to date these models cannot predict accurately the microstructural evolution in situations as complex as those experienced during rolling in a tandem mill. Furthermore, they often require knowledge of internal state variables that are difficult to measure, such as dislocation density during the thermomechanical process at all stages. To obtain accurate predictions in these complex situations, it is often necessary to take an empirical approach in which the input-output relationship is determined from the data without reference to any simplified physical model.

One such approach is neural network modeling. Neural networks offer a flexible approach to data modeling as they can provide an arbitrarily complex, nonlinear mapping between one or more inputs and the output. This mapping is parameter-

ized by a set of "weights," the optimum values of which are determined by training the network. Neural networks have been used with some success, particularly in materials science applications such as the prediction of weld toughness in steel[10] and of damage in forged composites.[11]

In this study a static Gaussian process (GP) model is applied to the problem of predicting crystallographic texture for PSC deformed and recrystallized AA3104 aluminum alloy.

2. Experimental Procedure

2.1 Plane Strain Compression Testing

The experimental program was undertaken using a computer-controlled plane strain compression (PSC) testing facility at Alcan International Limited's Banbury Research Laboratory, UK. This machine is built on a C-frame design with horizontally acting platens pressurized by a gas spring and with the main moving parts servohydraulically controlled. The key operating features are the two main cams, which travel at high velocity driven by the main hydraulic cylinder, which is powered by the stored energy from six accumulators. The cams are profiled to allow a constant strain rate on samples from 20 mm starting thickness down to 250 μm. The motion profile is communicated to the moving platen via tie bars that pull cam rollers onto the surface of the cams.[12] The machine is capable of compressing specimens under similar strain, strain rate, and temperature regimes as those typically found in industrial rolling operations. In particular, it is able to simulate hot tandem mill rolling, as it can perform multiple deformations in sequence on a single specimen based on a programmed profile.

2.2 Rolling Simulation Tests

A typical modern hot rolling mill comprises two stages: ingot breakdown rolling to produce a slab in several passes (typically 13 to 19) through a hot reversing mill, followed by rolling of the slab through the three or four stands of a hot tandem mill and coiling of the resulting strip. The tandem mill deformation simulations performed in this work were carried out using specimens machined from a sample of commercially hot rolled AA3104 alloy slab taken immediately prior to transfer to the hot tandem mill. The PSC specimens were 60 mm wide by 75 mm high and 10 mm thick with the shortest side parallel to the rolling direction. Thermocouple holes were drilled into each specimen to measure the temperature during the four stages of deformation. Each specimen was given a sequence of four deformations with the strains, strain rates, and temperatures selected to simulate a possible thermomechanical trajectory through a tandem mill. The trajectory for each PSC test was calculated using an in-house Alcan model for the hot tandem mill rolling process. Within this testing philosophy, the overall test matrix was defined to cover a wider range of deformation conditions than normally found in the commercial operation of a hot tandem mill. The ranges of strain, strain rate, and temperature covered by the test matrix for each of the four stages of deformation are given in Table 1. After the last deformation step the samples were quenched in a water bath. The specimens were then recrystallized by annealing at 350 °C for two hours, and the texture components were measured by conventional x-ray methods. A total of 52 samples were measured.

Table 1 Ranges of the Thermomechanical Parameters Used in the Experiment at Four Stages Representative of the Four Rolls

Deformation	Stage I	Stage II	Stage III	Stage IV
Strain	0.62-1.2	0.62-1.2	0.62-1.2	0.34-1.2
Strain rate	3.2-10.0/s	9.5-23.9/s	22.5-55.0/s	67.4-163.9/s
Temperature	397-550 °C	337-507 °C	276-462 °C	202-399 °C

3. Modeling Methods

Industrial processing of materials such as aluminum is complex. Although scientific investigations have helped greatly in understanding the underlying physical phenomena, there remain many problems where quantitative treatments are lacking. The slow rate of progress in predicting mechanical properties in commercial metal processes is due to the dependence of a large number of variables. Nevertheless, there are clear patterns that experienced metallurgists recognize and understand.

Neural network models are extremely useful in such circumstances, not only in the study of mechanical properties but wherever the complexity of the problem is overwhelming from a fundamental perspective and where simplification is unacceptable, as in the case of predicting the texture of rolled aluminum alloys.

Here we have used a specific model, namely the Gaussian process model, for predicting the annealed textures of PSC samples tested at different strain, strain rate, and temperature values. A brief description of the model is given in the paragraphs that follow. While being totally empirically based, such a model fulfills two important roles: first, it enables one to delineate better process control strategies for obtaining desirable microstructures, and second, it facilitates identifications of the key determinacies for guiding the subsequent development of physically based models.

3.1 The Gaussian Process Model

In modeling complex problems empirically, one does not know what the parameterized form of the input-output relationship should be *a priori*. The Gaussian process model is a way of avoiding having to explicitly parameterize this relationship by parameterizing a probability model over the data instead.[13,14]

The training data set, D, consists of a set of N inputs, X_N ($= \{x_1, x_2, \ldots, x_N\}$), and the corresponding N outputs, t_N. We are interested in interpolating these data using a model able to predict t at values of x that are not present in the training data. Generally, the measured values of t will contain noise, ν, so the model's prediction, $y(x)$, is related to the target output by $t = y(x) + \nu$.

A common approach to this noisy interpolation problem is to parameterize the function $y(x, w)$, where w is a set of parameters that are determined from the training data, using methods such as least-squares minimization of some cost function, E. This is the approach taken by feedforward multilayer perceptron neural networks, which provide a suitable framework for evaluating a nonlinear interpolating function (interpolant) of a set of training data. The parameters in these net-

works are presented by a set of weights: Training the neural network is the process of calculating the optimum weights by minimizing the cost function, $E(w)$.

Bayesian probabilistic data modeling is a robust approach to prediction problems and can be readily incorporated into the neural network approach. Rather than giving a single "optimum" prediction, Bayesian methods provide a probability distribution over the predicted value. This feature is often very important as it can be used to produce a characteristic error in the predictions, which represents the uncertainty arising from interpolating noisy data. The probability of the data given the weights, the likelihood, can be written $P(D|w,\beta) \propto \exp^{-\beta E}$. β is a so-called hyperparameter, which parameterizes the probability distribution and is related to the noise variance, v, in the target outputs. The maximum likelihood approach to training a neural network (minimizing E) is therefore equivalent to maximizing $P(D|w,\beta)$. However, we normally include an explicit prior on the weights to specify our belief concerning the distribution of the weights in the absence of any data. This can be written $P(w|\alpha)$, where α is another hyperparameter. This prior is often used to give preference to smoother interpolating functions rather than rapidly varying ones, which overfit the training data. This prior is particularly important when one tries to model sparse data sets, as it generally improves the reliability of predictions. We then apply Bayes' theorem to the prior and the likelihood to give the posterior probability distribution of the weights given the data

$$P(w|D,\alpha,\beta) = P(w|D,\beta)\, P(w|\alpha)/P(D) \qquad \text{(Eq 1)}$$

It is this quantity that we should maximize when training the neural network. (We can ignore the denominator, $P(D)$, when making predictions with a single model and data set, D.)

The Bayesian approach to predictions prescribes that one marginalizes (i.e., sums) over uncertain parameters. Ideally, therefore, one should integrate over all values of the weights rather than optimize them. We are interested in predicting a new value, t_{N+1}, given its corresponding input, x_{N+1}, and the set of training data, D. In terms of probability distributions we are interested in finding $P(t_{N+1}|x_{N+1},D,\alpha,\beta)$. This is obtained by integrating over all possible values of the weights:

$$P(t_{N+1}|x_{N+1},D,\alpha,\beta) = \int P(t_{N+1}|x_{N+1},D,w,\alpha,\beta)\, P(w|D,\alpha,\beta)\, dw \qquad \text{(Eq 2)}$$

The maximum of $P(t_{N+1}|x_{N+1},D,\alpha,\beta)$ yields the most probable prediction for t_{N+1}. The integration can be performed by Monte Carlo methods or by making simplifying assumptions about the form of $P(w|D,\alpha,\beta)$ in the equation.

Note that we are really interested in $P(t_{N+1}|x_{N+1},D)$ rather than $P(t_{N+1}|x_{N+1},D,\alpha,\beta)$. This is obtained from Eq 2 by also integrating over the hyperparameters α and β (although it is often adequate to optimize α and β). These hyperparameters are important because they control the complexity of the model. They are distinct from parameters (i.e., network weights), which parameterize the input-output mapping. One of the advantages of the Bayesian approach to data modeling is that it automatically embodies complexity control by means of these hyperparameters.

From the Bayesian perspective, we are interested only in $P(t_{N+1}|x_{N+1},D)$; we are not interested in the network weights themselves. Given that we should integrate over all weights, a preferable model is one that does not have such weights at all. The Gaussian process can be considered a neural network that we have integrated over all possible values of weights.

The Gaussian process approach to the prediction problem assumes that the joint probability distribution of any N output values, t_N, is an N-dimensional Gaussian distribution:

$$P(t_N|X_N,\Theta) = 1/Z\, \mathrm{Exp}[-1/2\, (t_N-\mu)^T C_N^{-1}(t_N-\mu)] \qquad \text{(Eq 3)}$$

This distribution is completely determined by the mean, μ, and the covariance matrix, C_N. The elements of C_N are given by the covariance function $C_{i,j} = C(x_i,x_j,\Theta)$, where x_i and x_j are any two inputs and Θ is a set of hyperparameters.

These hyperparameters explicitly parameterize a probability distribution over the input-output function rather the function itself. This is a probability distribution in an N-dimensional space. Training of the model is done by maximizing the probability of the hyperparameters. Once trained, the model predicts the most probable value of the output for a new set of inputs, together with a measure of the uncertainty of the prediction. The degree of correlation achieved by a given proximity of the input vectors is dictated by the hyperparameters. There is one of these hyperparameters for each input dimension. The relative size of the hyperparameters is a measure of relevance of each input dimension in determining the output. The model assesses the relevance of each input parameter for the prediction of the outputs from the associated values of the hyperparameters. This model prohibits the process of overfitting the training data, as models based on conventional neural network are prone to do, and finds the most generalized path to reach the output.[14] In the present case there are four loading steps, each corresponding to a stand on the tandem mill. If the process parameters are confined to strain, strain rate, and temperature, this gives 12 process variable for each sample.

4. Results and Discussion

The cube, S, Goss, Copper, and random texture components were measured for all of the 52 PSC specimens. In this exercise we have focused our analysis on predicting the cube and S texture components because S is the rolling texture and the cube is the main texture component that affects the earing.

The GP model was run using 52 data sets from 52 PSC test results with various combinations of variables and by changing the strain to true strain. Normally to assess the predictive capability of such a model, half the data are used for training and the other half for testing. However, our data set was too small to train the model on just half of the data and so we used a one-by-one analysis technique, which still has the virtue that the model has no prior knowledge of the outcome of the test trajectory. In this procedure, 51 data sets were used for training and the outcome of the 52nd input trajectory was predicted. This procedure was then repeated for each of the other 51 data sets. The results (predicted using 52 similar models) were plotted as the true values on the X-axis and the predicted data on the Y-axis. To compare the performance of the analyses for

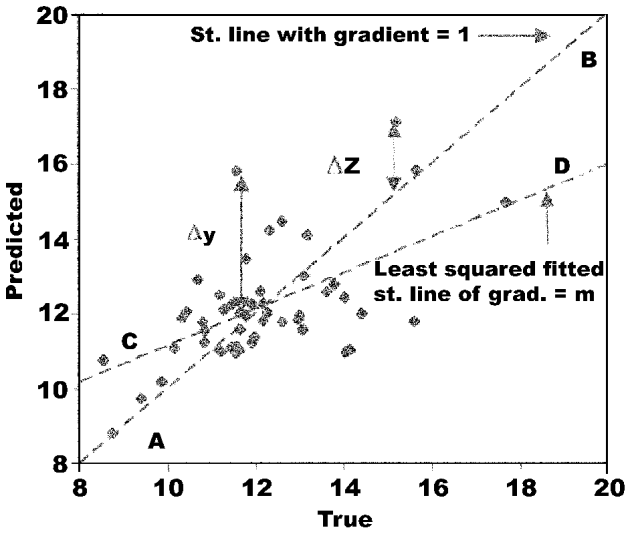

Fig. 1 Schematic depicting how the coefficients were measured

Table 2 Predictive Capability of Different GP Models as a Function of the Variables Included in Training

	Variables	M	R	r	⟨(ΔZ/σ)⟩
	%R, T, SR	0.48	0.023	0.018	1.13
	T, SR	0.57	0.019	0.015	0.97
Cube	*%R, T*	0.44	0.023	0.018	1.07
	TS, T, SR	0.45	0.024	0.019	1.24
	%R, SR	0.22	0.025	0.013	0.87
	%R, T, SR	0.82	0.016	0.015	0.93
	T, SR	0.67	0.220	0.019	0.95
S	*%R, T*	0.84	0.015	0.014	0.92
	TS, T, SR	0.82	0.016	0.015	0.93
	%R, SR	0.31	0.030	0.019	0.88

%R: Percentage reduction, *T:* Temperature, *SR:* Strain rate, *TS:* True strain

various combinations of the variables, we formulated various metrics for quantifying the goodness of the predictions. Figure 1 shows a typical true against predicted plot. On this figure the difference between the true value and the predicted value is denoted as ΔZ. This is also the distance of the ordinate of the predicted value from the straight line (*AB*) having a gradient equal to 1. In certain circumstances the predicted versus true plot exhibits a strong level of correlation, but the gradient is less than 1. Although such a relationship is not ideal, it could still be used to control the output given the inputs. The straight line *CD* is the least-squared fitted line of the predicted results having a gradient m. Δy is the distance of the ordinate of the predicted points from this line. The GP model also estimates the uncertainty (standard deviation σ) of the prediction of each datum. Consequently, the following metrics have been used are estimated based on the following mathematical formulae:

$$R = \frac{1}{N}\left\{\sum(\Delta Z)^2\right\}^{\frac{1}{2}} \quad \text{(Eq 4)}$$

$$r = \frac{1}{N}\left\{\sum(\Delta Y)^2\right\}^{\frac{1}{2}} \quad \text{(Eq 5)}$$

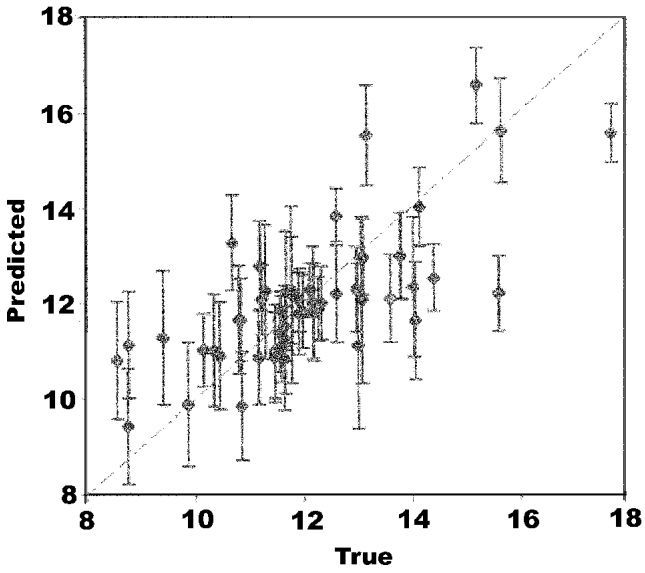

Fig. 2 Prediction of cube texture by GP model variables temperature and strain rate; the broken line is the 1:1 gradient.

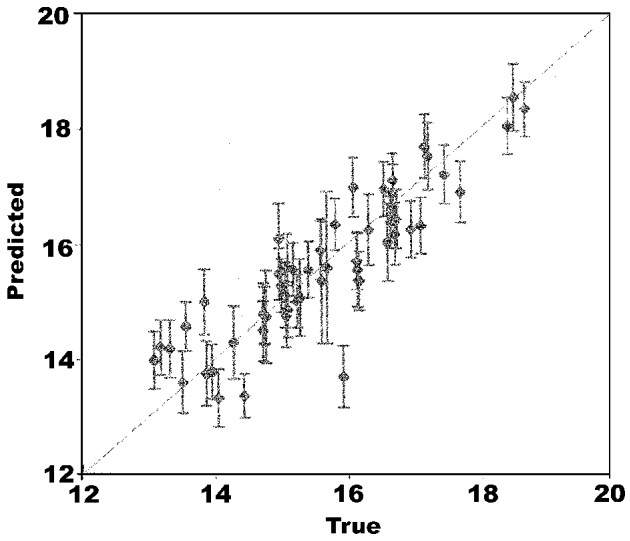

Fig. 3 Prediction of *S* texture by GP model variables percentage reduction and temperature; the broken line is the 1:1 gradient.

$$<(\Delta Z/\sigma)> = \frac{1}{N}\left\{\sum(\Delta Z/\sigma)^2\right\}^{\frac{1}{2}} \quad \text{(Eq 6)}$$

where *N* is the number of data points. In the calculations of *R* and *r,* the true values were normalized by the observed (measured) range so that the parameters vary between −0.5 (smallest observed value) to 0.5 (largest observed value). The predicted values were also normalized by the measured range. In principle, the ideal case is for m to be 1 and for the other coefficients to be 0. The coefficient *R* measures the root mean square (rms) error in the prediction and r measures the rms deviation of the predicted points from the least-squared fitted straight line. The rms $<(\Delta Z/\sigma)>$ is useful because large differences

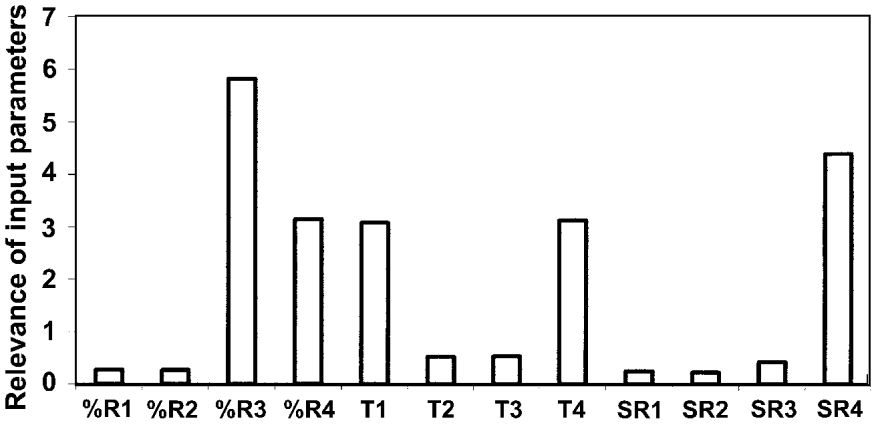

Fig. 4 Relevance of cube texture variables percentage reduction, temperature, and strain rate at four stages

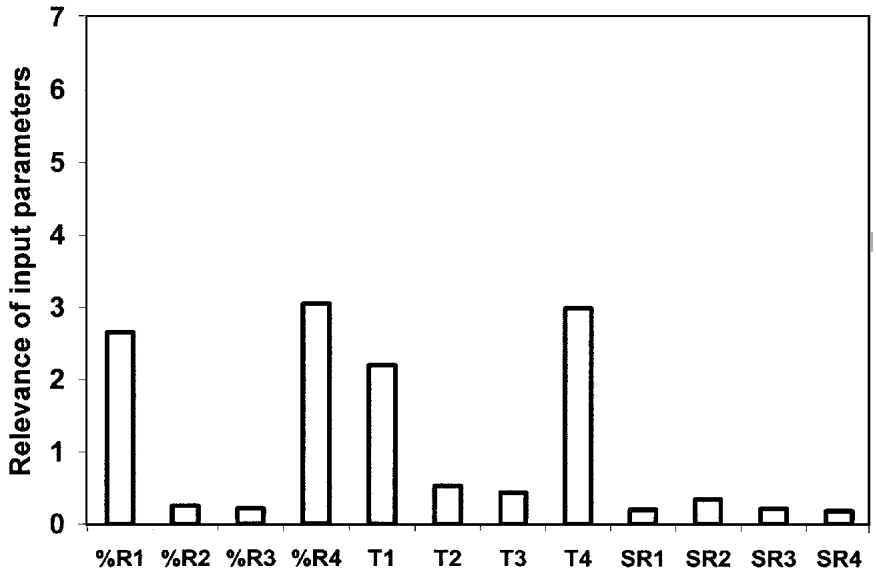

Fig. 5 Relevance of *S* texture variables percentage reduction, temperature, and train rate at four stages

between the model and the observed value are more serious when the model has inferred a high level of certainty in the prediction, while large differences between the model and the observed value would be expected if a large uncertainty were associated with the prediction of that point. The results of analyses are given in Table 2. Figure 2 and 3 show graphical representations for the prediction of the cube and *S* components of the recrystallized texture by the GP model, respectively. The uncertainty bars in these graphs are for one standard deviation, and thus approximately 66% of the predictions would be expected statistically to fall within them. Figure 4 and 5 show the relevance of the variables used in the prediction by the GP model for the cube and *S* texture components, respectively. A high relevance indicates that this parameter has a strong influence on the output.

The percentage reduction variables were converted to the more physically meaningful true strain values. However, this nonlinear transformation did not improve the prediction performance of the GP model for the cube and *S* texture components. This result also showed that percentage reduction is sufficient to represent the deformation strain. From the results shown in Table 2, it is clear from the GP model that temperature has a prime influence on texture. When this parameter was omitted, the prediction deteriorated dramatically. This effect also supports our understanding from the metallurgical point of view of the rolling process. Using the GP model for cube texture prediction, temperature and strain rate are the most important variables, whereas for *S* it is the temperature and the percentage reduction. This finding might suggest that the dynamic mechanisms for the formation of the two texture components are different. From the relevance plots we find that for the cube texture, temperatures in the first and the fourth stages of deformation, strain rate at the fourth stage of deformation, and strain at the third and fourth stages of deformations are all highly relevant. In the case of the *S* texture predictions, the temperatures and strain of the first and fourth stages of deformation are important, but strain rate at all stages has a low relevance.

5. Conclusions

The Gaussian process model has been used for the prediction of cube and *S* texture based on PSC test results. This model suggests that temperature is the prime variable that influences cube and *S* texture. Such models can be used to evaluate optimal strategies for the control of process parameters during four-stand tandem mill rolling.

Acknowledgments

The support of Alcan International is acknowledged. The authors are grateful for helpful advise from Dr. H.R. Shercliff.

References

1. R.A. Ricks: *Philos. Trans. R. Soc. London*, 1999, Ser.. A, *357*, pp. 1513-29.
2. I. Gutierrez and M. Fuentes: "Influence of the Microstructural Changes Occurring During Steady State Hot Deformation on Static Recrystallization Kinetics and Recrystallized Grain Size of Commercial Aluminum," *Recrystallization 90*., Ed., T Chandra TMS, 1990 Warrendale, PA pp. 807-12.
3. N. Raghunathan, M.A. Zaidi, and T. Sheppard: "Recrystallization Kinetics of Al—Mg Alloys AA 5056 and AA 5083 After Hot Deformation," *Mater. Sci. Technol*., 1986, *2*, pp. 938-45.
4. C.M. Sellers, A.M. Irisarri, and E.S. Puchi: "Recrystallization Characteristics of Aluminum—1% Magnesium Under Hot Working Conditions" in *Microstructural Control in Aluminum Alloys: Deformation, Recovery, and Recrystallization*, E. Henry Chia and H.J. McQeen, ed., The Metallurgical Society/AIME, Warrendale, PA, 1986, pp. 179-96.
5. H.E. Vatne, T. Furu, R. Orsund, and E. Nes: "Modeling Recrystallization After Hot Deformation of Aluminum," *Acta. Mater*., 1996, *44*, pp. 4463-73.
6. E.S. Puchi, J. Beynon, and C.M. Sellars: "Simulation of Hot Rolling Operations on Commercial Aluminum Alloys," *Proc. Int. Conf. On Physical Metallurgy of Thermomechanical Processing of Steels and Other Metals: THERMC '88*,. 1988, I. Tamura, ed., The Iron and Steel Institute of Japan, Tokyo, Japan, 1988, pp. 572-79.
7. P.D. Hodgson and R.K. Gibbs: "A Mathematical Model to Predict the Mechanical Properties of Hot Rolled C-Mn and Microalloyed Steels," *ISIJ Int*., 1992, *32*, pp. 1329-38.
8. R. Sandstrom and R. Lagneborg: "A Model for Static Recrystallization After Hot Deformation," Acta Metall., 1975, *23*, pp. 481-88.
9. T. Furu, H.R. Shercliff, G.J. Baxter, and C.M. Sellars: "The Influence of Transient Deformation Conditions on Recrystallization During Thermomechanical Processing of an Al-1% Mg Alloy," Acta Mater., 1999, *47*, pp. 2377-89.
10. H.K.D.H. Bhadeshia, D.J.C. Mackay, and L.E. Svensson: "Impact Toughness of C-Mn Steel Arc Welds—Bayesian Neural Network Analysis," *Mater. Sci. Technol*., 1995, *11*(10), pp. 1046-51.
11. S.M. Roberts, J. Kusiak, Y.L. Liu, A. Forcellese, and P.J. Withers: "Prediction of Damage Evolution in Forged Aluminum Metal Matrix Composites Using a Neural Network Approach," *J. Mater. Proc. Technol*., 1998, *80-81*, pp. 507-12.
12. G.J. Marshall: "Simulation of Commercial Hot Rolling by Laboratory Plane Strain Compression and its Application to Aluminum Industry Challenges." *Proc. of the Second Symposium of Hot Deformation of Aluminum Alloys II*, 1998, ed., T.R. Bieler, L.A. Lalli, and S.R. MacEwen, ed., The Minerals, Metals and Materials Society, Warrendale, PA, 1998, pp. 367-82.
13. D.J.C. Mackay: "Bayesian Interpolation," *Neural Comput*., 1992, *4*, pp. 415-47.
14. D.J.C. Mackay: "Probable Networks and Plausible Predictions-a Review of Practical Bayesian Methods for Supervised Neural Networks," *Network: Computat. Neural Syst*., 1995, *6*, pp. 469-505.