Modelling the Flow Stress of Alloy 316L using a Multi-Layered Feed Forward Neural Network with Bayesian Regularization

Olufunminiyi Abiri and Bhekisipho Twala
Institute of Intelligent Systems, University of Johannesburg, Auckland Park 2006, South Africa
olufunminiyi@uj.ac.za

Abstract. In this paper, a multilayer feedforward neural network with Bayesian regularization constitutive model is developed for alloy 316L during high strain rate and high temperature plastic deformation. The input variables are strain rate, temperature and strain while the output value is the flow stress of the material. The results show that the use of Bayesian regularized technique reduces the potential of overfitting and overtraining. The prediction quality of the model is thereby improved. The model predictions are in good agreement with experimental measurements. The measurement data used for the network training and model comparison were taken from relevant literature. The developed model is robust as it can be generalized to deformation conditions slightly below or above the training dataset.

1. Introduction
Material modelling is a challenging area in manufacturing processes such as machining. The large variation in strain, strain rates and temperature brings in complications leading to the problem of localized deformation in simulation [1]. In localized deformations, state variables such as strain, strain rate and temperature interactions are highly nonlinear and thus complete physical mechanisms are complex and difficult to understand [2]. Constitutive modelling of metals at high strain rates can be categorized into phenomenological, physical-based and artificial neural network (ANN), based models. Phenomenological models [3] are analytical model expressed in a convenient mathematical function. Physical-based models such as the dislocation density based model[4] combine and use physical and thermal deformation mechanisms such as thermodynamics and kinetics of slip. ANN modelling involves using an arbitrary number of computational processing units that learn from observed material data [5].

ANN is most suitable to capture the highly nonlinear relationships of the deformation behaviour of metals at high strain rate and temperature, as there is no need for a priori assumption of the mathematical model or a priori identifying material parameters that could be many. The material behaviour development is based on the experimental data available in training[6]. Also, material model based on ANN is very flexible and dynamic. The model can be extended within the training session as new experimental data about the behaviour of the material becomes available. This work aims to show the accuracy and robustness of ANN using Bayesian regularization in modelling alloy 316L during high strain rate and high temperature deformation. Instead of using gradient descent optimization methods, multi-layered feedforward neural network (MFNN) with Bayesian regularization[7] enables us to minimise overfitting and thus improve the generalization of the neural network model. Experimental data with both low and high strain rates for varying temperatures are utilized in the MFNN model. The data are taken from [8]. Strain ($\dot{\varepsilon}$), strain rate ($\dot{\varepsilon}$) and temperature
(T) are used as inputs of the MFNN, and the flow stress $\sigma$ is the output. Predicted flow stress results from the MFNN model are compared with the experimental measurements. The results indicate good agreement with measurements.

2. Multi-layered feedforward neural network
An ANN is a mathematical model involving a group of interconnected artificial neurons that tries to simulate the neural structure of the human brain. Artificial neural networks are trained against input and output data. This training process also called a learning process, is used to determine the neural network parameters for the model. An ANN is composed of an input layer, hidden layer or layers and an output layer. The layers are connected by independent processing units named neurons. The number of neurons in the input and output layers is determined by the number of input and output variables respectively. The number of hidden layer or layers depends on the level of complexity suitable for the modelling problem. The number of neurons in the hidden layer or layers is determined during the training process. An example of a neural network with one hidden layer is depicted in Figure 1. Data flow only in one direction through the network, from input to output. The mapping of the inputs to the output can be expressed in a MLP network written as

$$f(x; w, \omega) = g_2\left(\sum_{j=0}^{N} w_j g_1\left(\sum_{i=0}^{k} \omega_{ji} x_i\right)\right)$$

where $g_1(.)$ and $g_2(.)$ are transfer functions also known as activation functions. $w$ and $\omega$ are network weight parameters. The transfer function mathematically defines the non-linear relationship between inputs and output of a neuron in a neural network.

![Figure 1. Schematic illustration of neural network architecture](image)

The neural network parameters, $w$ and $\omega$, are unknown and they are determined by training the network. For the training set of data comprising a set of input $x_n$, $n = 1, ..., N$ and a corresponding set of target vectors $t_n$, the objective of ANN model is to minimize the sum of square error function derived from Equation (1). The error function is given as

$$E(\omega) = \frac{1}{2} \sum_{n=1}^{N} \left\| y(x_n; \omega) - t_n \right\|^2$$

where $t_n$ is the target output for the nth output neuron when the nth input pattern is presented to the network and $y(x_n, \omega)$ is the actual output of the nth output neuron when the input pattern is presented. This type of learning process is known as supervised learning in which every input pattern has its associated target output [9]. In ANN modelling, the goal is to find a vector of weights such that
Equation (2) takes its smallest value.

The artificial neural network makes use of an optimization algorithm to determine how good the error on the training input–output data set is minimized. It also determines the number of iteration steps. Overfitting defined as the effect of generalization error resulting from the overly complicated model in an optimization process may occur. Overfitting can be addressed by using Bayesian regularization methods [10,11]. The method automatically incorporates model uncertainty thereby imposing prior probability distribution on the model parameters. Using Bayesian regularization, \( E \) given by Equation (2) is modified by adding another term. The new error function, which can be called, augmented error function \( E_{aug}(\omega) \) becomes

\[
E_{aug}(\omega) = E(\omega) + \frac{1}{2} \mu \sum \alpha_{ij}^2
\]

where \( \mu \) and \( \alpha_{ij} \) are the regularization parameters known as the weight decay constant and weight connection from the node \( j \) to node \( i \) respectively. The additional terms lead to small values of the weight parameters thereby decreasing the tendency of the model to overfit the data.

3. The data set

The chemical composition of alloy 316L considered in this work is presented in

Table 1.

| C   | Si | Mn | P   | S   | Cr  | Mo  | Ni  | V   | N   |
|-----|----|----|-----|-----|-----|-----|-----|-----|-----|
| 0.009 | 0.31 | 1.71 | 0.031 | 0.023 | 16.86 | 2.04 | 10.25 | 0.048 | 0.040 |

The data set used in the development of the MFNN was taken from the flow stress measurement curves in Svoboda et al. [8]. The curves were obtained from Split Hopkinson Pressure Bar (SHPB) experimental study of the alloy. Experimental measurement \( T \) range from room temperature (RM) to 950 °C and \( \dot{\varepsilon} \) varies from 1000 s\(^{-1}\) up to 9000 s\(^{-1}\). The maximum \( \varepsilon \) obtained is about 60%. In this work, a total of 363 datasets were collected by digitizing the curves. This data set is used to train and test the developed MFNN model.

4. Model selection, experiment and simulation

The MFNN structure for the flow stress prediction for alloy 316L under high temperature and high strain rate deformation is depicted in Figure 2. Following the Kolmogorov theory[12], we are using a 3-7-1 network as shown in Figure 2. The network input layer consists of 3 neuron nodes \( \{\varepsilon, \dot{\varepsilon}, T\} \), the single hidden layer contains seven nodes and the output layer has only one node \( \sigma \). The activation function for the nodes in the hidden layer is a tangent sigmoid transfer function (TANSIG) while the output layer used the linear transfer function (PURELIN). The neural network package in the MATLAB software was used to develop and execute the model.
A learning rate of 0.01 was used. The number of epochs is 10000. An epoch can be defined as one presentation of all training examples to the network, followed by adjustment of the weights. The neural network was optimized according to the Bayesian regularization procedure in [13] using Equation (3) as the performance evaluation function. The Bayesian regularization method helps to determine the actual number of parameters, i.e. weights and bias, that are important to the NN model. All the 363 datasets from the flow stress curves is presented to the MFNN for training the network. The dataset composition is shown in Table 2. However, the dataset is pre-processed so that the input data will fall into the range $[-1,1]$. The MATLAB “mapminmax” is used for this purpose. This normalization procedure is used so that larger values of the input data do not overwhelm the smaller input data. This helps to reduce the network error. The output values were de-normalised. The neural network is trained on all of the data. The NN model was trained six times with different randomly generated initial parameters, and the best results are reported in this paper. After training, the MFNN model parameters are as in the appendix.

| $T$, °C | $\dot{e}$, s$^{-1}$ | Total number in the dataset |
|--------|-----------------|--------------------------|
| Room temperature | 2850 | 38 |
| Room temperature | 8345 | 37 |
| 200 | 2930 | 34 |
| 400 | 2965 | 48 |
| 700 | 2965 | 44 |
| 500 | 8670 | 43 |
| 500 | 1020 | 25 |
| 900 | 1075 | 44 |
| 900 | 8810 | 50 |

5. Results and discussions

Figure 3 shows the mean square error (MSE) curve after the MFNN training with the dataset. The MSE decreases as a function of the number of epochs. At convergence, the error is 0.000301, which is not far from the ideal value of zero. The linear regression analysis, in terms of the target and network’s output is shown in Figure 4. The R-value is 0.999. In the figure, the target and the network output exhibit good correlation, and they fall almost on the same line with the angle of the passing of the output line 45°. This illustrates that the network has been trained properly with no occurrence of overfitting. The predictive ability of the MFNN model for the training dataset is shown in Figure 5. The plot provides excellent matching of the measurement. The predictive accuracy and robustness of the MFNN model are further shown by comparing with the experimental measurement for the alloy at the various condition of in [14]. This comparison is shown in Figure 6. The experimental measurement in the figure is obtained by digitizing the appropriate curves in [14]. The MFNN model
shows good agreement with the measurement over the full range of data and also compared very well with the dislocation density based physical model results in [14]. The MFNN generalise very well for deformation conditions that are slightly below of above the training dataset.

Figure 3. Training curve for the MFNN.

Figure 4. Linear regression analysis for the trained MFNN.

Figure 5. MFNN output versus measured flow stress for the training dataset.

Figure 6. Comparison of MFNN predicted and measured flow stress for alloy 316L at various conditions of temperature and strain rates.

6. Conclusions
The flow stress of alloy 316L at high temperature and high strain rate deformation has been modelled using an artificial neural network with Bayesian regularization, MFNN. The input level for the neural network consists of strain rate, temperature and strain and the output level consist of the flow stress. High strain rate and high temperature deformation of 316L leads to localized deformation that is highly nonlinear and a challenge for modelling and simulation. From this work, the following conclusion can be drawn:

i. Using ANN with Bayesian regularization affords us the opportunity of combining nonlinear functions in parallel for the modelling of the alloy.
ii. The predicted results from the MFNN model are in good agreement with experimental measurement and also consistent with the physical based results in [14].

iii. The MFNN is robust as it can generalize to deformation condition slightly below or above the training dataset and therefore able to capture the different physical mechanisms of the alloy during high strain rate and high temperature deformation.

iv. The model can be improved as more experimental measurement becomes available for the alloy.

Thus, the MFNN model developed in this paper is viable for use in the simulation manufacturing processes, particularly machining.

Acknowledgements
The authors gratefully acknowledge the financial support of the University of Johannesburg via the Global Excellence and Stature (GES) fellowship granted to the first author.

References

[1] Lindgren L, Svoboda A, Wedberg D, Lundblad M. Towards predictive simulations of machining. Comptes Rendus Mécanique 2016;344:284-295.
[2] Vaz Jr M, Owen D, Kalhori V, Lundblad M, Lindgren L. Modelling and simulation of machining processes. Archives of computational methods in engineering 2007;14(2):173-204.
[3] Johnson GR. Response of boron carbide subjected to large strains, high strain rates, and high pressures. J Appl Phys 1999;85(12):8060-8073.
[4] Wedberg D, Svoboda A, Lindgren L. Modelling high strain rate phenomena in metal cutting simulation. Modell Simul Mater Sci Eng 2012;20(8):085006.
[5] Sheikh-Ahmad J, Twomey J. ANN constitutive model for high strain-rate deformation of Al 7075-T6. J Mater Process Technol 2007;186(1):339-345.
[6] Ghaboussi J, Garrett Jr J, Wu X. Knowledge-based modeling of material behavior with neural networks. J Eng Mech 1991;117(1):132-153.
[7] Sizo DI, Twala B and Marwala T. Predictive modeling for default risk using a multilayered feedforward neural network with Bayesian regularization. Neural Networks (IJCNN), The 2013 International Joint Conference on: IEEE; 2013.
[8] Svoboda A, Wedberg D, Lindgren L. Simulation of metal cutting using a physically based plasticity model. Modell Simul Mater Sci Eng 2010;18(7):075005.
[9] Bishop CM. Neural networks for pattern recognition. : Oxford university press; 1995.
[10] MacKay DJ. A practical Bayesian framework for backpropagation networks. Neural Comput 1992;4(3):448-472.
[11] Neal RM. Bayesian learning for neural networks. PhD Thesis 1995;University of Toronto.
[12] Hecht-Nielsen R. Theory of the backpropagation neural network. Neural Networks 1988,1(Supplement-1):445-448.
[13] Gauss-Newton approximation to Bayesian learning. Neural Networks, 1997., International Conference on: IEEE; 1997.
[14] Wedberg D, Lindgren L. Modelling flow stress of AISI 316L at high strain rates. Mech Mater 2015 12;91, Part 1:194-207.