

Article



# Hybrid Machine Learning Optimization Approach to Predict Hot Deformation Behavior of Medium Carbon Steel Material

## Mohanraj Murugesan, Muhammad Sajjad and Dong Won Jung \*

Department of Mechanical Engineering, Jeju National University, Jeju-Do 63243, Korea; mohanaero32@jejunu.ac.kr or mohanaero45@gmail.com (M.M.); msajjadtheone@gmail.com (M.S.)

\* Correspondence: jdwcheju@jejunu.ac.kr

Received: 11 November 2019; Accepted: 3 December 2019; Published: 6 December 2019



Abstract: The isothermal tensile test of medium carbon steel material was conducted at deformation temperatures varying from 650 to 950 °C with an interval of 100 °C and strain rates ranging from 0.05 to  $1.0 \text{ s}^{-1}$ . In addition, the scanning electron microscopy (SEM) procedures were exploited to study about the surface morphology of medium carbon steel material. Using the experimental data, the artificial neural network (ANN) model with a back-propagation (BP) algorithm was proposed to predict the hot deformation behavior of medium carbon steel material. For model training and testing purpose, the variables such as deformation temperature, strain rate, and strain data were considered as inputs and the flow stress data were used as targets. Before running the neural network, the test data were normalized to effectively run the problem and after solving the problem, the obtained results were again converted in order to achieve the actual data. According to the predicted results, the coefficient of determination  $(R^2)$  and the average absolute relative error between the predicted flow stress and the experimental data were determined as 0.999 and 1.335%, respectively. For improving the model predictability, the constrained nonlinear function based optimization procedures was adopted to obtain the best candidate selections of weights and biases. By evaluating each test conditions, it was found that the average absolute relative error based on the optimized ANN-BP model varied from 0.728% to 1.775%. Overall, the trained ANN-BP models proved to be much more efficient and accurate by means of flow stress prediction against the experimental data for all the tested conditions. These optimized results displayed that an ANN-BP model is more accurate for flow stress prediction than that of the conventional flow stress models.

**Keywords:** isothermal tensile test; medium carbon steel; surface morphology; neural network; back-propagation; flow stress

## 1. Introduction

Medium carbon steel materials are generally employed for a wide range of engineering applications due to their vital mechanical properties such as wear resistance and weldability [1]. It is obvious that most of the automobile components are manufactured by causing the plastic deformation in the work material and also the manufactured components experience the plastic deformation in real time applications. To explain the material mechanics at thermal processing, the constitutive relationship, which describes the flow stress by relating it to strain, strain rate, and deformation temperature, is commonly employed [2–4]. Therefore, it is important to understand the work material deformation behavior in hot working involving high deformation temperatures and strain rate conditions. Moreover, the constitutive mathematical equations, which developed from experiment observations, will be helpful for the numerical modeling of the metal forming process

like forging, ring rolling and stamping process; also the computational simulations can reduce time and cost required for modeling the experiments at development stages. To describe the material behavior, several number of mathematical models are proposed, but the available flow stress models such as Johnson–Cook (JC), modified Johnson–Cook (MJC), modified Zerilli–Armstrong (MZA), and Arrhenius-type constitutive (AC) models displayed a low accuracy on the flow stress prediction in some tested materials [5–9]. The researchers found that it is because most of the existed constitutive models are constructed from the regression models; it is because the factors associated with the flow stress are highly nonlinear and very complex. In recent years there has been considerable attention on the neural network model to develop the flow stress models during hot working conditions. Based on our previous [8,9] research experiences, the aim of our research work was further extended to propose a mathematical model using neural networks for evaluating the work material flow behavior.

In order to describe the material behavior under hot deformations, for past decades, a considerable number of research articles have been published using existed flow stress models such as JC, MJC, MZA, and AC models. Li et al. [10] researched about T24 steel material flow behavior under elevated temperatures and high strain rates using JC and MJC models. They stated that the proposed MJC model results agreed well with experimental observations rather than the original JC model. He et al. [11] studied the effects of strain, strain rate and temperature in 20CrMo alloy steel material to predict the hot deformation behavior using JC, MJC, and AC models. Among the developed models, MJC and AC models showed better agreement with the experimental data, whereas the traditional JC model had worked well only under reference test conditions. Samantaray et al. [12] did comparative study on JC, MZA, and AC models to describe the flow behavior of modified 9Cr-1Mo steel material and they observed that the conventional JC model displayed inadequate results against the real data, but other models provided a good description of flow behavior. Li et al. [13] did the isothermal uniaxial tensile tests under elevated temperatures (20–900 °C) and high strain rates (0.01–10.0 s<sup>-1</sup>) for four kinds of boron steel B1500HS material and investigated the problem using modified AC and MJC models. They concluded that from comparison with a computational data, the constitutive equations provided better correlations against the experimental measurements.

Li et al. [14] built comparative study on MJC, MZA, and AC models to explain hot deformation behavior in 28CrMnMoV steel under hot working conditions. They concluded that MJC and MZA models showed good agreement against the experimental data except at low strain rates, however, overall strain-compensated AC displayed better correlation with the real data than others. Lei et al. [15] conducted constitutive analysis to describe high temperature flow behavior of 3Cr-1Si-1Ni ultra-high strength steel and proposed an AC model that accounted strain compensation showed much capability in describing the material behavior at hot working conditions. He et al. [16] verified existing flow stress models capability to capture the flow behavior in Ti2AlNb-based alloys and the constitutive equation displayed a more precise description against the experimental data. Yang et al. [17] and Li et al. [18] researched the advantage of using the strain-compensated AC equation to establish the flow stress model for super-austenitic and Nb-contained 316LN stainless steel materials during hot working, respectively. They claimed that the modified AC equation provided a high accuracy in flow stress predictions. Aforementioned flow stress models require much time to do fitting equations for the estimation of material constants, model development, evaluation and verification against the experimental observations. Therefore, constructing flow stress model in terms of the analytical model that provides an accurate prediction outcome based on the relationship between the experimental input conditions and the observations will be more helpful.

In recent years, for this purpose, the artificial neural network (ANN) model has been employed as the model uses mathematical formulations to construct a brain nervous system operation based on relationships exist between inputs and outputs. Moreover, a countable number of research articles were published with respect to this topic for the application of flow stress prediction at hot working conditions [19–23]. Ji et al. [24] and Peng et al. [25] studied about developing constitutive relationship at elevated temperatures and strain rates test conditions using an AC and ANN models in Aermet100

steel and as-cast Ti60 titanium alloy materials, respectively. They drew the conclusion that the back-propagation (BP) ANN model can accurately predict the desired data in the entire working domain. In addition, the flow stress estimation in a wide range of test conditions associating phase transformations in as-cast Ti–6Al–2Zr–1Mo–1V alloy was investigated by Quan et al. [26]. They pointed out that the ANN model can forecast material flow behavior including the metallurgical phenomenon, and also stated that the model has a capability to capture complex behavior even outside of the test conditions. Ashtiani et al. [27] and Stendal et al. [28] employed both phenomenological and ANN models to predict high-temperature deformation behavior in AlCuMgPb alloy and titanium aluminide alloy (TNM-B1) materials. They also identified that the well trained ANN model can be able to make accurate predictions than the tested phenomenological equations. Han et al. [29] proposed a model from an AC and an ANN models for as-cast 904L austenitic stainless steel to predict the material deformation behavior and results proved that the optimized ANN model has ability to

hot working conditions. Although there is rapidly growing literature on an ANN model, there are only limited articles to discuss improving an ANN model accuracy. Huang et al. [30] proposed a modified ANN-BP based on genetic algorithm (GA) to predict the material behavior in aluminum alloy. They outlined that an ANN-GA model displayed a more efficient and accurate prediction capacity for the entire test process. This present research aims to achieve an optimized ANN-BP model for computing the flow stress of AISI-1045 medium carbon steel at hot working conditions. In order to identify the suitability of an ANN-BP model, the experiment was conducted under wide range of deformation temperatures (650–950  $^{\circ}$ C) and strain rates (0.05–1.0 s<sup>-1</sup>); thereafter, the flow stress data was collected and arranged eventually. In the network modeling process, the parameters such as strain, strain rate and temperature were considered as input variables and the flow stress was used as output variable, respectively. For obtaining optimized random weights and biases, the optimization procedures (OP) based on constrained nonlinear function were modeled and utilised for training, validating and testing an ANN-BP model to develop the material flow stress model. Eventually, the computed results from ANN-BP and ANN-BP/OP models are tested against experimental measurements and constructed model verification's are discussed using both graphical and numerical validations.

capture the compressive behavior at high deformation temperatures. From literature survey, recent studies outcome indicates the importance of ANN model to characterize the material flow behavior at

#### 2. Material and Experimental Procedures

In this research work, AISI-1045 medium carbon steel material was used and the chemical composition (in wt.%) of the work material is summarized in Table 1. The isothermal tensile tests were modeled at deformation temperatures (*T*) ranging from 650 °C to 950 °C and strain rates ( $\dot{\epsilon}$ ) varying from 0.05 s<sup>-1</sup> to 1.0 s<sup>-1</sup>, respectively. The samples were prepared according to ASTM-E8M-subsize standard and all samples were thoroughly checked for dimensional accuracy; and then used for conducting experiments. The experimental procedures used in this present work to characterize medium carbon steel material behavior as follows: at first, the specimens were produced from the water-jet cutting process and three samples were used for each test condition. Secondly, the load-displacement curves were transformed into true stress ( $\sigma$ )–strain ( $\epsilon$ ) (SS) curves and then obtained SS curves were averaged lately as depicted in Figure 1a–d. From Figure 1a–d, the plastic regions are separated in order to construct the ANN-BP model for establishing the constitutive relationship of work material. The experimental set-up used in our research work is more detailed in our previous research papers [8,9].

Table 1. Chemical composition of AISI-1045 medium carbon steel (in wt.%)

С	Fe	Mn	Р	S	
0.42-0.50	98.51-98.98	0.60-0.90	$\leq 0.04$	$\leq 0.05$	



**Figure 1.** True stress-strain data obtained from hot tensile tests at various temperatures under different strain rates.

For evaluating micro-structure evolution in AISI-1045 medium carbon steel material, the field emission scanning electron microscopy (FESEM) (MIRA3 TESCAN, secondary electron detector, Jeju National University, Jeju-si, Korea) along with the energy dispersive X-ray spectroscopy (EDS) mapping setup was utilised in this research work. Using the test setup, the deformation temperature (850 °C, 950 °C) and strain rate (0.05–0.5 s<sup>-1</sup>) dependent surface morphology, thickness and elemental identification analysis were observed at the fractured surface for various magnifications as illustrated in Figure 2. From Figure 2a,e,i, for strain rates at deformation temperature (850 °C), it can be seen that specimen growth and nucleation was found to be coarse at low strain rate and as strain rate kept increasing, the nucleation and growth transformation was noticed to be finer, uniform, and homogeneous at a 50  $\mu$ m scale. As seen in Figure 2b,f,j, the FESEM images are presented with micro and nanopores, a highly interconnected porous structure observed at 20 µm scale. However, observation at a higher magnification scale level (5 µm) revealed that a porous and interconnected structure was found to be discontinuous and irregular as shown in Figure 2c,g,k. In addition, Figure 2i,k comparisons against Figure 2m,n show that due to temperature changes, a macro porous structure was formed and furthermore, we observed the apparent growth, nucleation of nanoneedles, and reduction of pore size. Moreover, the magnified portion (at 5 µm scale) of the sample image confirms the formation of micro-fibrils and moderate growth of nanoneedles. The reason behind this clear difference in microstructure is possibly due to slow and strong self-association of grains at low temperature. Besides, at magnification 100 µm scale, for test conditions (850 °C) and (950 °C) at strain

rate ( $0.5 \text{ s}^{-1}$ ), the EDS mapping analysis of test specimens proves the presence of iron (Fe) and carbon (C) elements throughout the scanned surface. The microstructure images of fractured specimens at 500 µm scale, as shown in Figure 2d,h,l,o, confirms the reduction of specimen thickness at the fracture location at tested conditions. The inset images show the rough surface morphology at fractured surface (5 µm scale).



**Figure 2.** Micro-structure mapping images of AISI-1045 medium carbon steel material at deformation temperature (850 °C). (**a**–**d**)  $0.05 \text{ s}^{-1}$ ; (**e**–**h**)  $0.1 \text{ s}^{-1}$ ; (**i**–**l**)  $0.5 \text{ s}^{-1}$ ; (**m**–**o**) 950 °C/0.5 s<sup>-1</sup>; observation by FESEM and EDS on various magnifications.

## 3. Artificial Neural Network Approach

Machine learning (ML) problems are classified into three sections: supervised, unsupervised, and reinforcement learning based on the data availability. In supervised learning, the supervisor already knows the working conditions and its responses in a restricted environment. So by selecting proper supervised learning algorithms, experimental inputs can be mapped into the outputs that

6 of 19

have the best solution. On the other hand, with the assumption of much more probability and the maximum model metrics to its goal, other two problems are devised for achieving the best candidate solution. An ANN model provides a novel approach to the model prediction of material processing as it does not need to postulate a mathematical model or identify its parameters. Moreover, ANN shows a good ability to predict the flow behavior because it is particularly suitable for treating complex and non-linear relationships between the responses of deformation behaviors of the materials under elevated temperatures and high strain rates [31–35].

#### 3.1. Flow Stress Modeling of AISI-1045 Steel Using an ANN with Back-Propagation Algorithm

In this research work, a multi-layer feed forward ANN model with supervised learning procedure, BP algorithm for training, was employed to construct the functional relationship among input and output variables for predicting flow stress at hot working conditions as shown in Figure 3. As can be seen in Figure 3, there were three input variables: strain ( $\varepsilon$ ), strain rate ( $\dot{\varepsilon}$ ), and deformation temperature (*T*), and one output variable, stress ( $\sigma$ ), in the neural network design. The network training began with initialization of random weights and biases, and then the weight was adjusted based on the prediction error between computed and experimental observations using BP learning algorithm. In detail, the process involved two steps: forward and backward pass. At first, the model inputs were fed into the network with casual weights and arrived at the output layer with the approximate solution. Thereafter, in a backward pass, the output from the forward pass was compared against the real data, and then the weights and biases were altered iteratively to minimize the mean square error (MSE) (Equation (1)) with the help of BP learning algorithm and the process was continuous for each input–output pair throughout the modeling process [36].

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (actual_i - predicted_i)^2$$
(1)

In Equation (1), n is the total number of samples used for training the network. After training the network with each input–output pair, the trained network model c ouldbe tested using unknown input–output sets for evaluating the performance of the proposed network model.



**Figure 3.** Back-propagation artificial neural network (BP-ANN) architecture for flow stress prediction (supervised learning).

It is obvious that there is no proper ground rule or fixed procedures for selecting initial random weights, the size of training data set, total number of neurons in the hidden layer, learning algorithm (weight adjustment) and transfer functions in order to construct the ANN-BP model. In this research work, from literature survey, the detailed procedures to adopt the hidden layer with the optimum number of neurons, number of samples, learning algorithms and activation functions are presented.

Carpenter et al. [37] suggested that in the hidden layer, number of neurons in terms of minimum counts can be explained using the following expressions:

$$HN = IN + 1$$

$$HN = 3 + 1 = 4$$
(2)

whereas, in Equation (2), HN and IN are number of neurons in the hidden layer and number of variables in the input layer, respectively. Accordingly, for this research work, in the hidden layer, number of neurons varied from 2 to 30 to confirm the capability of Equation (2).

Further, for training the neural network model, a minimum number of required sample data are estimated as follows [38,39]:

$$NT = HN * (IN + 1) + NO * (HN + 1)$$
  
NT = 4 \* (3 + 1) + 1 \* (4 + 1) = 21 (3)

In Equation (3), NO and NT are number of variables in the output layer and number of training data, respectively. It is important to mention here that sample counts are essential in order to effectively train the network as it has a huge influence on network generations for producing close predictions. Using Equation (3), the minimum number of samples for one test condition were estimated as 21 and for all conditions, the total number of samples were calculated as 336. In addition, for improving performance of network generations and also for capturing a wide range of variations, the sample counts were increased from 336 to 384 and kept constant throughout the modeling process. The significant logic behind 384 data selection: in this research work, 16 sets of test conditions were investigated thorough the real time experiments for capturing material deformation behavior. From SS curves as shown in Figure 1, the plastic range in terms of strain ( $\varepsilon$ ) was identified as 0.02 to 0.25. Thus, in order to eliminate the inconsistent results from under and over fitting of training data, the data are evenly considered with the strain increment of 0.01 and eventually, 24 samples were accounted for each test condition to construct the neural network model.

Before the network training process, entire input and output variables are normalized in order to obtain a usable form for the network using Equation (4).

$$X_{\rm N} = \frac{X - 0.95 X_{\rm min}}{1.05 X_{\rm max} - 0.95 X_{\rm min}} \tag{4}$$

where *X* is the measured experimental data,  $X_{\min}$  and  $X_{\max}$  are the minimum and maximum values of chosen actual data such as stress ( $\sigma$ ), strain ( $\varepsilon$ ), strain rate ( $\dot{\varepsilon}$ ), and deformation temperature (*T*), respectively, and  $X_N$  is the normalized data. The experimental values are normalized between more than 0 and less than 0.95, because in the end points, the transfer functions showed a slow learning rate behavior while training the network model [40]. Likewise, data samples (100%) are randomly partitioned into three sets as training set (70%), validation set (15%) and test set (15%) as listed in Table 2 in order to perform the network training process. Training and validation sets are used for training the network and monitoring the training process, respectively, whereas the test set is used to examine the performance of trained network in untrained samples.

Number of samples Input layer Hidden layer functions Number of neurons Output layer Output layer function network type net algorithm Training functions Learning function	384 data (268 (Training) + 58 (validation) + 58 (Testing)) three variables LOGSIG and TANSIG two $\leq$ HNs $\leq$ 30 one variable Purelin multi-layer feed-forward back-propagation Trainbr and Trainlm LEARNGDM
Learning function	LEARNGDM
Performance function	MSE

Table 2. Network model instructions used for constructing flow stress model.

The selection of activation functions also influences the performance of the neural network in terms of capturing the function approximations among input and output variables. For choosing transfer function in the hidden layer, two most widely used functions such as tan-sigmoid (Equation (5)) and log-sigmoid (Equation (6)), are adopted, because from Figure 1a-d, the variations ( $\sigma$ ) are noticed to vary continuously but not linearly as changing input variables (T,  $\dot{\epsilon}$ ,  $\epsilon$ ). The reason for choosing sigmoid functions is that it always produces an "S" shape output as shown in Figure 4e,f; the shape is tends to be linear in the middle and nonlinear towards the boundaries. From Figure 4e,f, the transfer function, log-sigmoid, was noticed to produce only positive values, which is not suitable if the network expects to return negative values during the training process, whereas tan-sigmoid delivered values from positive to negative, which is suitable in any cases. However, some trail and error calculation procedures were required to select the activation function, in which the network always contributed the optimal solutions. In addition, among the available training functions, trainbr (Bayesian regularization) and trainlm (Levenberg-Marquardt) functions were picked based on their capability to learn to map inputs to outputs within given data-set. For the output layer, the transfer function was directly selected as purelin (linear function) because the problem assumed to be linear in the output layer as the model output was proportional to the total weighted inputs.

tansigmoid function : 
$$a = \frac{2}{1 + \exp(-2n)} - 1$$
 (5)

logsigmoid function : 
$$a = \frac{1}{1 + \exp(-n)}$$
 (6)

Using an ANN-BP model informations summarized in Table 2, the network models were trained, validated, and tested against the experimental observations. Thereafter, the computed results of MSE and coefficient of determination ( $R^2$ ) are plotted against neurons as depicted in Figure 4a–d. Figure 4a–d distinctly show that there is little difference between activation functions in terms of MSE and  $R^2$  values. But somehow, there are considerable deviations among the training functions and trainbr function is found to be the best selection for training the network as it displays steady improvement in terms of error decrements along with the highest correlation ( $R^2$ ) value. As expected based on Equation (2), from Figure 4a–d, it is confirmed that the prediction error is significantly higher value up to four neurons. As listed in Table 3, It is obvious that more number of neurons lead to higher accuracy, but however, after 18 neurons (trainbr), error sums are fluctuating in a random manner. This fluctuation conveys that in order to control over-fitting with unknown points, the size of neurons should be limited to acceptable margin. Therefore, considering the network model complexity, the error differences are inspected closely from 4 to 30 neurons and identified that the predicted results are reasonably accurate when the network contains eight neurons in the hidden layer.



**Figure 4.** Verification procedures for obtaining accurate ANN-BP models. (**a**–**d**) Comparison of training functions; (**e**,**f**): Activation functions.

MSE				MSE					
Neurons	Neurons TANSIG		LOGSIG		Neurons	TANSIG		LOGSIG	
	Trainbr	Trainlm	Trainbr	Trainlm	-	Trainbr	Trainlm	Trainbr	Trainlm
2	311.190	314.301	311.187	565.457	18	0.730	12.535	0.705	13.121
4	99.431	42.762	43.316	42.352	20	2.139	14.458	1.874	13.095
6	14.149	15.586	11.079	87.875	22	0.555	6.252	0.486	1.187
8	13.219	20.846	5.767	9.807	24	1.202	9.919	1.383	1.713
10	3.417	10.782	4.128	28.448	26	1.141	23.703	1.078	2.008
12	2.562	4.871	2.361	13.318	28	0.776	2.097	0.642	18.675
14	1.405	17.683	1.504	5.243	30	1.144	9.715	0.930	6.473
16	1.082	8.037	1.066	3.351					

Table 3. Effect of number of neurons in hidden layer on performance of ANN-BP model.

Ultimately, an ANN-BP model consists of one hidden layer with eight neurons, trainbr algorithm as training function, learngdm algorithm as a learning function, and one output layer is chosen for predicting the flow stress at hot working conditions. Moreover, the network model performance also depends on learning parameters, such as the number of training epochs and the performance goal, etc. But in this work, the number of epochs, the learning rate, and the error threshold were fixed to a certain level based on the literature survey as 1000, 0.05 and  $1 \times 10^{-06}$ , respectively. The MSE quantity between actual and predicted data was recorded during network model training and using minimized or converged MSE value, the best models were obtained for both activation functions. Now the trained ANN-BP model should be verified to make a confirmation that the model implementation was done correctly. It is obvious that, in machine learning process, the model performance evaluation and its interpret-ability are the vital procedures for pointing out a solid conclusion about the proposed ANN-BP model. Here, we exploited some helpful procedures such as quantification and visualization

methods to comment about the developed flow stress model accuracy in an explicit manner. In addition, there are a few other possible evaluation checks available like examining the proposed model in unknown points, technically called as untrained samples, to monitor the model predictability. These interpretation choices are discussed here in detail, particularly, by both numerical and graphical validations. Moreover, the evaluation techniques presented in this research work is significantly sufficient to confirm the model capability, because the prediction outcomes are always tested against experimental observations. For quantification purpose, three kinds of statistical parameters such as  $R^2$ , average absolute relative error (AARE), and root mean square error (RMSE) are utilized.  $R^2$  explains the prediction strength against experimental observations in terms of a quantity ranging from 0 to 1, whereas AARE is employed to quantify the prediction error at overall test conditions against the experimental data. RMSE is a statistical measure of differences between values predicted by a proposed model and the values actually measured from the experiments [41].

$$R^{2} = 1 - \frac{\sum_{i=1}^{n} (\sigma_{e}^{i} - \sigma_{ANN}^{i})^{2}}{\sum_{i=1}^{n} (\sigma_{e}^{i} - \bar{\sigma}_{e})^{2}},$$
(7)

$$AARE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\sigma_{e}^{i} - \sigma_{ANN}^{i}}{\sigma_{e}^{i}} \right| \times 100\%,$$
(8)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (actual_i - predicted_i)^2}{n}}$$
(9)

where  $\sigma_e$ ,  $\sigma_{ANN}$ ,  $\bar{\sigma}_e$  are the experimental data, the predicted flow stress from an ANN-BP model, and the mean flow stress, respectively, and *n* is the total number of data points.

Using Equations (7) and (8), the population parameters are computed for each test conditions and summarized in Table 4. It is clearly seen that both transfer functions in the hidden layer displayed a significantly better outcome. For tansig activation function,  $R^2$  and AARE were estimated as 0.9980 and 1.3348%, respectively, and whereas for logsig activation function,  $R^2$  and AARE were determined as 0.9991 and 1.8059%, respectively. But, in test conditions, 850 °C and 950 °C, the prediction error was found to be higher, but it was significantly acceptable as the error value was close to 2.2%. Apart from the numerical validation, the graphical validation against the experimental observations was plotted as depicted in Figure 5 for tansig activation function to prove an ANN-BP model capability. Figure 5 displays that the model predictions were scattered closer to the experimental data and it confirms that the ANN-BP model can provide accurate representation of material flow behavior under hot deformation conditions. Moreover, the plastic-instability phenomenon also tended to be captured more effectively than that of available traditional flow stress models [8,9].

 Table 4. Computed statistical parameters from an ANN-BP model.

ANN transfer function	Test Conditions		$R^2$	Overall- <i>R</i> <sup>2</sup>	AARE (%)	Overall-AARE (%)
	$0.05 - 1.0 \text{ s}^{-1}$	923 K 1023 K	0.9918 0 9990	0.9980	1.6397 1.4028	1.8059
TANSIG		1023 K 1123 K 1223 K	0.9995		2.1722	
		923 K	0.9971		0.8637	
LOGSIG	$0.05 - 1.0 \text{ s}^{-1}$	1023 K 1123 K 1223 K	0.9996 0.9997 0.9998	0.9991	0.8927 1.4321 2.1507	1.3348



**Figure 5.** Comparison between experimental and predicted flow stress data using BP-ANN model with TANSIG as a hidden layer activation function .

#### 3.2. Optimization Procedures for Obtaining the Best Trained ANN-BP Model

In the neural network model, the training process is carried out using an iterative process, which means in each step the model is updated with small weights and biases, for finding an optimum set of weights and biases to improve the model performance. A general approach for solving the neural network problem is to restart the training process multiple times with different random initial weights and biases, and allow the searching algorithm to find distinct candidates for the best trained ANN-BP model. This process is usually called multiple restarts. In this research work, the multiple restarts process was modeled by employing hybrid optimization procedures for training a network model in terms of adjusting weights and biases to predict the flow stress of medium carbon steel material under hot deformations as shown in Figure 6. The nonlinear programming function, find minimum of constrained nonlinear multivariable function (fmincon), was utilized considering the interior-point (IP) algorithm to minimize AARE between an ANN-BP model and the desired flow stress data; the bounds constrained optimization procedures exploited in this work is also depicted in Figure 7. The IP algorithm was selected due to its advantage in finding the minimum of a function within the presence of bounds constraints. Moreover, the benefits of exploiting this fmincon function rather than GA is that the computational time to solve the problem can be minimized without compromising the accuracy of

results eventually. Wang et al. [42] stated that in their research, the optimization problem based on fmincon function immensely reduced the calculation time related with GA, and also they pointed out that the obtained results was found to be reasonable against the actual results. In general, it is difficult to mention whether using wide range of bounds are valid or not at the first place. Therefore, at start of the optimization process, the problem was tested with a small range of bounds and then increased a little wider for allowing the process to be sampled extensively before selecting a valid candidate for a better solution. The limits of bounds constraints are selected from the trail experience for solving the optimization problem; the general form of optimization procedures are expressed below:





Figure 6. BP-ANN model with an OP for flow stress prediction.



Figure 7. Flow chart of optimization procedures for minimizing network model prediction error.

At start of the optimization process, the initial points are given randomly in order to begin optimization. At next step, the fmincon function automatically alters the  $x_0$  points strictly between the bounds constraints. Subsequently, the optimization algorithm searches through a bounds space of feasible values for the neural network model for a set of weights and biases that results in good performance on the model outcome. The optimization problem for activation functions is terminated when there is no improvement in the solution towards feasible directions and also the constraint tolerance is satisfied within the specified margin. The best candidate solutions for tansigmoid function in the hidden layer are obtained when the iterations and the function counts are 14 and 183, respectively, whereas for logsigmoid function, the numbers are computed as 5 and 71, respectively. The optimum solutions of AARE with transfer functions, tansig, and logsig, are achieved as 1.123% and 1.502%, respectively. The optimal results computed from the proposed ANN-BP/OP model are tabulated in Table 5. Tables 4 and 5 are strong evident that the prediction error obtained from the optimized network model varies from 0.728% to 1.775%, whereas for the basic network model, errors are ranging from 0.8637% to 2.172%, which states that the optimized ANN-BP model can correlate the material flow behavior more effectively than the conventional network model. In addition, there was no considerable differences between tansigmoid and logsigmoid functions with regard to the prediction error, but somehow, the optimum network model with tansigmoid function looks a little significant as far as reduction in the prediction error is concerned.

ANN transfer function	Test Conditions		$R^2$	Overall-R <sup>2</sup>	AARE (%)	Overall-AARE (%)
TANSIG	$0.05 - 1.0 \text{ s}^{-1}$	923 K 1023 K 1123 K 1223 K	0.9940 0.9997 0.9998 0.9999	0.9989	1.1582 0.7282 1.0089 1.5963	1.1229
LOGSIG	$0.05 - 1.0 \text{ s}^{-1}$	923 K 1023 K 1123 K 1223 K	0.9960 0.9992 0.9996 0.9999	0.9988	1.0972 1.3804 1.7752 1.7541	1.5017

Table 5. Computed statistical parameters from an optimized ANN-BP model.

As can be seen in Figure 8, most of the predicted data points from the optimized ANN-BP model were close to the experimental measurements and this finding confirms the capability of the optimized flow stress model compared to the conventional network model. The correlation between experimental observations and the predicted is interesting because the computed data points almost followed the same trend line along the desired values as illustrated in Figure 8. Moreover, Figure 9a shows that the proposed model displayed a better correlation with respect to the measured data along with a better correlation coefficient  $R^2$  value at 0.9989. In addition, the statistical measurements,  $R^2$  and AARE, were estimated for each test condition using the proposed model as summarized in Table 5 and likewise, it displays the excellent prediction ability of the proposed network model. Figure 9b,c displays the random distribution of residuals with respect to zero error line; also, from the histogram plot (inset images), the distribution of residuals was noticed to be random and the probability distribution was found to be normal inside the working space. Furthermore, Figure 9c conveys that even after the optimization process, the residual plot showed a fairly random pattern, which indicates that the proposed model provided a modest fit to the desired data. In addition, in order to clearly depict the model performance, the prediction error comparison using an ANN-BP and an ANN-BP/OP was modeled at different deformation temperatures and strain rates as shown in Figure 10a.



**Figure 8.** Comparison between experimental and predicted flow stress data using BP-ANN/OP model with TANSIG as a hidden layer activation function.



**Figure 9.** Graphical validation of network models. (**a**) Correlation plot; (**b**) Residual plot of an ANN-BP model; (**c**) Residual plot of an optimized ANN-BP model. (inset histogram plots).

5

2.3 2.5

5

4.5

3.5

3

2

1.5

Error (%)





Figure 10. Graphical validation of proposed flow stress models. (a) BP-ANN and BP-ANN/OP models comparison at various strain rates; (b) Prediction error comparison for proposed phenomenological, BP-ANN and BP-ANN/OP models.

According to our findings, the developed ANN models can be effectively applied to predict the material deformation behavior of medium carbon steel. Also the prediction error variations occurred in the traditional flow stress models [8,9], as shown in Figure 10b, that introduced by the plastic instability phenomenon can be eliminated. Overall, the presented discussion implies that the proposed ANN-BP model has more impact to deal with a nonlinear experimental data than that of the conventional flow stress models in order to approximate the constitutive relationship of medium carbon steel at hot working conditions.

#### 4. Conclusions

The material flow behavior of AISI-1045 medium carbon steel was developed using an artificial neural network model with a back-propagation learning algorithm for a wide range of deformation temperatures (650–950 °C) and strain rates (0.05–1.0 s<sup>-1</sup>). At first, without optimization procedures the model was developed and the predicted results from the proposed network model displayed a good agreement with the experimental measurements in terms of better correlation and low prediction error quantities. Subsequently, a hybrid algorithm was utilized for obtaining the best trained ANN-BP model to predict the flow stress of medium carbon steel material. The optimization procedures used to find the model parameters such as weights and biases that result in minimum prediction error when evaluating accuracy of an ANN-BP model against the experimental observation. From obtained results, it was found that an optimized BP-ANN with tansigmoid activation function displayed the much more accurate prediction capability to describe the material hot deformation behavior throughout the entire tested conditions. Moreover, the statistical measurements such as  $R^2$  and AARE, were calculated as 0.9989 and 1.1229%, respectively. In terms of the statistical parameters outcome, the optimized ANN-BP model was found to track the material behavior more effectively and also the better flow stress prediction for an entire temperature and strain rate range was achieved. Moreover, there were no mathematical model assumptions and physical insight needed to develop an ANN-BP model and these kind of procedures make it more effective to predict the material behavior than the conventional constitutive equations. The proposed models were more significant inside the tested conditions, and by contrast, the model was quiet weak outside the tested conditions.

Author Contributions: In this research, the parts of conceptualization, methodology, numerical modeling, programming, investigation, validation, and original draft preparation were done by the authors M.M. and M.S., and the supervision was carried out by D.W.J.

**Funding:** This research was supported by the 2019 scientific promotion program funded by Jeju National University.

**Conflicts of Interest:** The authors declare no conflict of interest.

### Abbreviations

The following abbreviations are used in this manuscript:

JC	Johnson–Cook
MJC	modified Johnson–Cook
MZA	modified Zerilli-Armstrong
AC	Arrhenius-type constitutive
ANN	artificial neural network
BP	back-propagation
SS	stress–strain
FE	finite element
SEM	scanning electron microscopy
FESEM	field emission scanning electron microscopy
EDS	energy dispersive X-ray spectroscopy
ε	strain (mm/mm)
Ė	strain rate (s <sup><math>-1</math></sup> )
Т	deformation temperature (°C)
σ	flow stress (MPa)
ML	Machine learning
п	number of samples
HN	number of neurons in hidden layer
IN	number of variables in input layer
NO	number of variables in output layer
NT	number of training data
$X_{\rm N}$	normalized data
Χ	measurements from experiment
X <sub>min</sub>	minimum value of experimental data
X <sub>max</sub>	maximum value of experimental data
TANSIG	Tan-Sigmoid
LOGSIG	Log-Sigmoid
$R^2$	coefficient of determination
RMSE	root mean square error
MSE	mean square error
AARE	average absolute relative error
OP	optimization procedures
fmincon	find minimum of constrained nonlinear multivariable function
IP	interior-point
GA	genetic algorithm
TOL	Tolerance
$W_{ij}$	network weights
$b_i$	network biases
IW	weights in hidden layer
LW	weights in output layer
b1	biases in hidden layer
b2	biases in output layer
trainbr	Bayesian regularization
trainlm	Levenberg-Marquardt
learngdm	Gradient descent with momentum weight and bias learning function
~	

#### References

- 1. Rhim, S.-H.; Oh, S.-I. Prediction of serrated chip formation in metal cutting process with new flow stress model for AISI 1045 steel. *J. Mater. Process. Technol.* **2006**, *171*, 417–422. [CrossRef]
- He Y.-B.; Pan, Q.-L.; Chen, Q.; Zhang, Z.-Y.; Liu, X.-Y.; Li, W.-B. Modeling of strain hardening and dynamic recrystallization of ZK60 magnesium alloy during hot deformation. *Trans. Nonferrous Met. Soc.* 2012, 22, 246–254. [CrossRef]
- Zhan, H.; Wang, G.; Kent, D.; Dargusch, M. Constitutive modelling of the flow behaviour of a β titanium alloy at high strain rates and elevated temperatures using the Johnson–Cook and modified Zerilli–Armstrong models. *Mater. Sci. Eng. A* 2016, 612, 71–79. [CrossRef]
- 4. Paturi, U.M.R.; Narala, S.K.R.; Pundir, R.S. Constitutive flow stress formulation, model validation and FE cutting simulation for AA7075-T6 aluminum alloy. *Mater. Sci. Eng. A* **2014**, *605*, 176–185. [CrossRef]
- Chen, G.; Ren, C.; Ke, Z.; Li, J.; Yang, X. Modeling of flow behavior for 7050-T7451 aluminum alloy considering microstructural evolution over a wide range of strain rates. *Mech. Mater.* 2016, 95, 146–157. [CrossRef]
- 6. Lee, K.; Murugesan, M.; Lee, S.M.; Kang, B.S. A comparative study on Arrhenius-type constitutive models with regression methods. *Trans. Mater. Proc.* **2017**, *26*, 18–27. [CrossRef]
- 7. He, Z.; Wang, Z.; Lin, Y.; Fan, X. Hot Deformation Behavior of a 2024 Aluminum Alloy Sheet and its Modeling by Fields-Backofen Model Considering Strain Rate Evolution. *Metals* **2019**, *9*, 243. [CrossRef]
- 8. Murugesan, M.; Jung, D.W. Johnson Cook Material and Failure Model Parameters Estimation of AISI-1045 Medium Carbon Steel for Metal Forming Applications. *Materials* **2019**, *12*, 609. [CrossRef]
- 9. Murugesan, M.; Jung, D.W. Two flow stress models for describing hot deformation behavior of AISI-1045 medium carbon steel at elevated temperatures. *Heliyon* **2019**, *5*, 1347. [CrossRef]
- 10. Li, H.Y.; Wang, X.F.; Duan, J.Y.; Liu, J.J. A modified Johnson Cook model for elevated temperature flow behavior of T24 steel. *Mater. Sci. Eng. A* 2013, 577, 138–146. [CrossRef]
- 11. He, A.; Xie, G.; Zhang, H.; Wang, X. A comparative study on Johnson–Cook, modified Johnson–Cook and Arrhenius-type constitutive models to predict the high temperature flow stress in 20CrMo alloy steel. *Mater. Des.* **2013**, *52*, 677–685. [CrossRef]
- 12. Samantaray, D.; Mandal, S.; Bhaduri, A.K. A comparative study on Johnson Cook, modified Zerilli–Armstrong and Arrhenius-type constitutive models to predict elevated temperature flow behavior in modified 9Cr–1Mo steel. *Comput. Mater. Sci.* **2009**, *47*, 568–576. [CrossRef]
- 13. Li, H.; He, L.; Zhao, G.; Zhang, L. Constitutive relationships of hot stamping boron steel B1500HS based on the modified Arrhenius and Johnson–Cook model. *Mater. Sci. Eng. A* **2013**, *580*, 330–348. [CrossRef]
- 14. Li, H.-Y.; Li, Y.-H.; Wang, X.-F.; Liu, J.-J.; Wu, Y. A comparative study on modified Johnson Cook, modified Zerilli–Armstrong and Arrhenius-type constitutive models to predict the hot deformation behavior in 28CrMnMoV steel. *Mater. Des.* **2013**, *49*, 493–501. [CrossRef]
- 15. Lei, B.; Chen, G.; Liu, K.; Wang, X.; Jiang, X.; Pan, J.; Shi, Q. Constitutive Analysis on High-Temperature Flow Behavior of 3Cr-1Si-1Ni Ultra-High Strength Steel for Modeling of Flow Stress. *Metals* **2019**, *9*, 42. [CrossRef]
- 16. He, Z.; Wang, Z.; Lin, P. A Comparative Study on Arrhenius and Johnson–Cook Constitutive Models for High-Temperature Deformation of Ti2AlNb-Based Alloys. *Metals* **2019**, *9*, 123. [CrossRef]
- Yang, L.-C.; Pan, Y.-T.; Chen, I.-G.; Lin, D.-Y. Constitutive Relationship Modeling and Characterization of Flow Behavior under Hot Working for Fe–Cr–Ni–W–Cu–Co Super-Austenitic Stainless Steel. *Metals* 2015, *5*, 1717–1731. [CrossRef]
- Li, J.; Liu, J. Strain Compensation Constitutive Model and Parameter Optimization for Nb-Contained 316LN. *Metals* 2019, 9, 212. [CrossRef]
- Zhu, Y.; Zeng, W.; Sun, Y.; Feng, F.; Zhou, Y. Artificial neural network approach to predict the flow stress in the isothermal compression of as-cast TC21 titanium alloy. *Comput. Mater. Sci.* 2011, 50, 1785–1790. [CrossRef]
- 20. Guo, L.-F.; Li, B.-C.; Zhang, Z.-M. Constitutive relationship model of TC21 alloy based on artificial neural network. *Trans. Nonferrous Met. Soc. China* **2013**, *23*, 1761–1765. [CrossRef]
- 21. Bobbili, R.; Madhu, V.; Gogia, A.K. Neural network modeling to evaluate the dynamic flow stress of high strength armor steels under high strain rate compression. *Def. Technol.* **2014**, *10*, 334–342. [CrossRef]

- 22. Xiao, X.; Liu, G.Q.; Hu, B.F.; Zheng, X.; Wang, L.N.; Chen, S.J.; Ullah, A. A comparative study on Arrhenius-type constitutive equations and artificial neural network model to predict high-temperature deformation behavior in 12Cr3WV steel. *Comput. Mater. Sci.* **2011**, *62*, 227–234. [CrossRef]
- 23. Li, H.-Y.; Wei, D.-D.; Li, Y.-H.; Wang, X.-F. Application of artificial neural network and constitutive equations to describe the hot compressive behavior of 28CrMnMoV steel. *Comput. Mater. Sci.* **2012**, *35*, 557–562. [CrossRef]
- 24. Li, G.J.F.; Li, Q.; Li, H.; Li, Z. A comparative study on Arrhenius-type constitutive model and artificial neural network model to predict high-temperature deformation behaviour in Aermet100 steel. *Mater. Sci. Eng. A* **2011**, *528*, 4774–4782.
- Peng, W.; Zeng, W.; Wang, Q.; Yu, H. Comparative study on constitutive relationship of as-cast Ti60 titanium alloy during hot deformation based on Arrhenius-type and artificial neural network models. *Mater. Des.* 2013, 51, 95–104. [CrossRef]
- 26. Quan, G.-Z.; Lv, W.-Q.; Mao, Y.-P.; Zhang, Y.-W.; Zhou, J. Prediction of flow stress in a wide temperature range involving phase transformation for as-cast Ti–6Al–2Zr–1Mo–1V alloy by artificial neural network. *Mater. Des.* **2013**, *50*, 51–61. [CrossRef]
- Ashtiani, H.R.R.; Shahsavari, P. A comparative study on the phenomenological and artificial neural network models to predict hot deformation behavior of AlCuMgPb alloy. J. Alloys Compd. 2016, 687, 263–273. [CrossRef]
- 28. Stendal, J.A.; Bambach, M.; Eisentraut, M.; Sizova, I.; Weiß, S. Applying Machine Learning to the Phenomenological Flow Stress Modeling of TNM-B1. *Materials* **2019**, *9*, 220. [CrossRef]
- Han, Y.; Qiao, G.; Sun, J.; Zou, D. A comparative study on constitutive relationship of as-cast 904L austenitic stainless steel during hot deformation based on Arrhenius-type and artificial neural network models. *Comput. Mater. Sci.* 2013, 67, 93–103. [CrossRef]
- 30. Huang, C.; Jia, X.; Zhang, Z. A Modified Back Propagation Artificial Neural Network Model Based on Genetic Algorithm to Predict the Flow Behavior of 5754 Aluminum Alloy. *Materials* **2018**, *11*, 855. [CrossRef]
- 31. Rath, S.; Talukdar, P.; Singh, A.P. Application of Artificial Neural Network for Flow Stress Modelling of Steel. *Am. J. Neural Netw. Appl.* **2017**, *3*, 36–39. [CrossRef]
- 32. Rao, S.G.R.N.; Nandy, T.K.; Bhattacharjee, A. Artificial neural network approach for prediction of Titanium alloy Stress-Strain Curve. *Procedia Eng.* **2012**, *38*, 3709–3714.
- Wu, R.H.; Liu, J.T.; Chnag, H.B.; Hsu, T.Y.; Ruan, X.Y. Predication of the flow stress of 0.4C-1.9Cr-1.5Mn-1.0Ni-0.2Mo steel during hot deformation. *J. Mater. Process. Technol.* 2001, 116, 211–218. [CrossRef]
- 34. Senthilkumar, V.; Balaji, A.; Arulkirubakaran, D. Application of constitutive and neural network models for prediction of high temperature flow behavior of Al/Mg based nanocomposite. *Comput. Mater. Sci.* **2013**, *23*, 1737–1750. [CrossRef]
- 35. Lin, Y.C.; Chen, X.-M. A critical review of experimental results and constitutive descriptions for metals and alloys in hot working. *Mater. Des.* **2011**, *32*, 1733–1759. [CrossRef]
- 36. Aghasafari, P.; Salmi, H.A.M. Artificial Neural Network Modeling of Flow Stress in Hot Rolling. *Comput. Mater. Sci.* 2013, 54, 872–879. [CrossRef]
- 37. Carpenter, W.C.; Barthelemy, J.F. Common Misconceptions about Neural Networks as Approximators. *ASCE J. Comput. Civ. Eng.* **1994**, *8*, 345–358. [CrossRef]
- 38. Carpenter, W.C.; Hoffman, M.E. Training backprop neural networks. J. AI Expert 1995, 10, 30–33.
- Oreta, A.W.C.; Kawashima, K. Neural network modeling of confined compressive strength. J. Struct Eng. 2003, 129, 554–561. [CrossRef]
- 40. Razavi, S.V.; EI-Shafie, A.H.; Mohammadi, P. Artificial neural networks for mechanical strength prediction of lightweight mortar. *Sci. Res. Essays* **2011**, *6*, 3406–3417. [CrossRef]

- 41. Murugesan, M.; Kang, B.S.; Lee, K. Multi-Objective Design Optimization of Composite Stiffened Panel Using Response Surface Methodology. *J. Compos. Res.* **2015**, *28*, 297–310. [CrossRef]
- 42. Chuan, W.; Lei, Y.; Jianguo, Z. Study on Optimization of Radiological Worker Allocation Problem Based on Nonlinear Programming Function-fmincon. In Proceedings of the International Conference on Mechatronics and Automation, Tianjin, China, 3–6 August 2014.



 $\odot$  2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).