Comparing the linear and logarithm normalized extreme learning machine in flow curve modeling of magnetorheological fluid

Irfan Bahiuddin¹, Abdul Y Abd Fatah², Saiful A Mazlan³, Mohd I Shapiai³, Fitrian Imaduddin⁵, Ubaidillah⁶, Dewi Utami⁷, Mohd N Muhtaruddin⁸

¹³⁷Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
²⁸Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Kuala Lumpur, Indonesia
⁵⁶Mechanical Engineering Department, Faculty of Engineering, Universitas Sebelas Maret, Surakarta, Indonesia

ABSTRACT

The extreme learning machine (ELM) plays an important role to predict magnetorheological (MR) fluid behavior and to reduce the computational fluid dynamics (CFD) calculation cost while simulating the MR fluid flow of an MR actuator. This paper presents a logarithm normalized method to enhance the prediction of ELM of the flow curve representing the MR fluid rheological properties. MRC C1L was used to test the performance of the proposed method, and different activation functions of ELMs were chosen to be the neural networks setting. The Normalized Root Mean Square Error (NRMSE) was selected as the indicator of the ELM prediction accuracy. NRMSE of the proposed method is found to improve the model accuracy up to 77.10% for the prediction or testing case while comparing with the linear normalized ELM.

Keywords: Extreme learning machine, Magnetorheological fluid, Neural networks, Normalized method, Shear stress

1. INTRODUCTION

A smart material called magnetorheological (MR) fluid is known for its capability to change the viscosity as the effect of magnetic fields exposure. The changes are reversible that happen less than 10 milliseconds. The highly quick responsive nature of the materials has attracted the interest of various researchers to apply at various devices such as dampers [1]-[3], brakes [4], clutches [5] and others [6]-[8]. In each application, the behavior of the MR fluid needs to be considered, especially in design stages. The usual important parameters or variables of the MR fluid rheological properties for design and simulation are shear stress, shear rate, viscosity and yield stress [9], [10]. These variables can be obtained from the characterization process using a rheometer. One of the common measurement results from the rheometer is called flow curve. Flow curve is a data set containing shear stress as a function of shear rate. This data set can be derived to obtain yield stress, plastic viscosity and also to identify whether the fluid is Newtonian or non-Newtonian, having shear thickening or thinning effects, and other relevant characteristics [11], [12].

Rheological models are the usual methods to predict and duplicate the MR fluid behavior in terms of the flow curve data. Rheological models include Bingham Plastic, Papanastasiou, Herschel Bulkley, and biviscous model [13]-[15]. Bingham Plastic is the most widely known method to obtain yield stress and plastic viscosity from a flow curve because of its simplicity [16], [17]. Another model is Herschel Bulkley that tried
to overcome the disadvantage of Bingham plastic model, which is the high accuracy for a narrow shear rate range [18], [19]. In general, all models can only be applicable to one value of the magnetic fields [20], [21]. If one wants to cover other magnetic fields, a polynomial model of yield stress needs to be applied [22], [23]. The limitation can be solved using machine learning, such as Artificial Neural Networks (ANN), Support Vector Machine (SVM) [24] and Extreme Learning Machine (ELM) [25]. ANN and SVM have been known for their limitations, which are slower training time and have more possibility to be trapped at local minima or extrema [26]. Extreme Learning Machine (ELM) is known for its solution in terms of the quicker training, more accuracy value and highest possibility to gain accepted generalization than the classic methods. The existing method to predict MR fluid behavior using ELM has been proposed in [20] to predict various rheological parameters. Despite the high accuracy at some values and ranges, the prediction needs to be improved. The possible improvement is the activation function, normalization method and the pre-processing of the inputs. Meanwhile, the activation function effects were investigated in [20], [25]. However, the normalization effect at the ELM-based rheological model is only discussed separately and not comprehensively. While linear normalization has been utilized in [25], [27], [28], logarithm normalization has been introduced in [20]. Since these previous works focus only on the introduction of the extreme learning machine platform for various important parameters prediction, the logarithm normalization influence on the model performance has not been discussed in detail.

Therefore, this paper presents a comparative study between the linear and logarithm normalized extreme learning machine in the flow curve modeling of the magnetorheological fluid. The normalization is applied to the inputs of the extreme learning machine (ELM). The model output is shear stress while the inputs are shear rate and magnetic field. The general steps of the modeling method, the normalization method, the data collection and the simulation set up are described in section 2. The results on the accuracy and computational burden are then presented and discussed in section 3.

2. RESEARCH METHOD

2.1. Modeling method

The usual rheological models have a shear rate as input and shear stress as output. Meanwhile, the ELM based-rheological model consists of the same input and output with the additional input of magnetic fields to predict the shear stress at various magnetic fields. The ELM based-rheological model is calculated using the structure of Single-hidden Layer Feedforward Neural Networks (SLFNs). In general, the output model or the predicted shear stress for each sample (i) set can be expressed as in Equation (1) or (2).

$$\sum_{j=1}^{L} \beta_{j} g(w_{j}, b_{j}, x_{i}) = o_{i}, i = 1, \ldots, N$$

(1)

$$H \beta = O$$

(2)

Output ($o_{i}$) is a function of weighting outputs ($\beta_{j}$) and hidden node output ($g(w_{j}, b_{j}, x_{i})$). $g(w_{j}, b_{j}, x_{i})$ can also be called Activation function as a function of bias ($b_{j}$) on j-th hidden node where $j = 1, \ldots, L$, and i-th inputs ($x_{i}$) where $i = 1, \ldots, N$. Weighting input ($w_{j}$) on each i-th input. Meanwhile, the hidden node outputs for complete training datasets can be denoted as H. The detail matrix of H is expressed by Equation (3). The predicted shear stress matrix for all samples is denoted as O that can be expressed in Equation (4).

$$H = 
\begin{bmatrix}
g(w_{1}, b_{1}, x_{1}) & \ldots & g(w_{L}, b_{L}, x_{1}) \\
\vdots & \ddots & \vdots \\
g(w_{1}, b_{1}, x_{N}) & \ldots & g(w_{L}, b_{L}, x_{N})
\end{bmatrix}_{N \times L}$$

(3)

$$O = 
\begin{bmatrix}
o_{1}^{T} \\
o_{L}^{T}
\end{bmatrix}_{N \times 1}$$

(4)

The activation function can utilize any continuous piecewise functions. The examples of the existing continuous functions are:
a. Hard Limit
\[
g(w, b, x) = \begin{cases} 1, & \text{if } w \cdot x - b \geq 0 \\ 0, & \text{if } w \cdot x - b < 0 \end{cases}
\]  
(5)

b. Sigmoid
\[
g(w, b, x) = \frac{1}{1 + \exp(-(w \cdot x + b))}
\]  
(6)

c. Sine
\[
g(w, b, x) = \sin(w \cdot x + b)
\]  
(7)

d. Triangular basis function
\[
g(w, b, x) = \text{if } -1 \leq n \leq 1; = 0, \text{otherwise}
\]  
(8)

e. Radial Basis Function
\[
g(w, b, x) = \exp(-n^2)
\]  
(9)

The error formulation between the experimental and simulation results is formulated using Equation (10) where T matrix is the measured shear stress of all samples as described by Equation (11)

\[
e = ||O - T||
\]  
(10)

\[
T = \begin{bmatrix} t_{11}^T \\ \vdots \\ t_{L1}^T \end{bmatrix}_{N \times m}
\]  
(11)

If e is near zero, we obtain \(O = T\). Replacing \(O\) in Equation (2) with \(T\) yields Equation (12).

\[
H\beta = T
\]  
(12)

ELM algorithm has been proven to minimize e and \(||\beta||\) [26]. The SLFNs model is trained using the ELM algorithm. In general, the ELM algorithm can be described within three steps, which are

a. assigning the \(w_j\) and \(b_j\) randomly using any continuous activation function,

b. calculating \(\beta\) by solving Eq. (8) where \(H^\dagger\) is defined as the Moore–Penrose generalized inverse [34]. \(H^\dagger\) was calculated using Singular Value Decomposition Method (SVD).

\[
\beta = H^\dagger T
\]  
(13)

2.2. Normalization method

The inputs usually need to be normalized before inputting the value of the Neural Networks. Therefore, all inputs can be inputted into the model with the same scale, between 0 and 1 or -1 and 1 depending on the designer decision. The normalization can be applied in the input or output variables as shown in Figure 1. The simplest form of normalization method is linear as expressed in the following equations,

\[
x_{\text{norm}} = \frac{x_e - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} (x_{\text{max}} - x_{\text{min}})
\]  
(14)

where \(x_{\text{norm}}, x_e, x_{\text{max}}, x_{\text{min}}\) are the normalized, experimental data, maximum, and minimum inputs, respectively.
Meanwhile, the normalized logarithm is expressed in Equation (15). Although normalized logarithm in the work of ANN for aircraft cabin has been proven to obtain better performance compared to the linear one [29], the application in MR fluid should be further investigated to gain more comprehensive analysis at various activation functions. In this work, the normalization is applied in the inputs, either one of or both shear rate and magnetic fields.

\[
x_{1,i} = \frac{\log(x_e) - \log(x_{min})}{\log(x_{max}) - \log(x_{min})}
\]

**2.3. Data Collection**

The flow curve is obtained based on the rotational test on a parallel plate rheometer (Anton Paar) as reported in [25]. The tested MR fluid is manufactured by CK, South Korea called MRC C1L with the properties as described in [25]. The applied magnetic fields are 830 mT, 770 mT, 420 mT, 310 mT, 200 mT, and 0 mT. The observation is carried out from 0.01 to 1000 s\(^{-1}\). The shear rate is taken based on the setting of a logarithmic ramp in Rheocompass, the graphical user interface of the parallel plate rheometer by Anton Paar. The obtained data are shown in Figure 2 that contain flow curve data about 539 measurement points. A number of data with the same applied magnetic field are called data-set. Thus, there are 7 datasets. The data is divided into training data and prediction. For MRC C1L, training data is used for obtaining the model. It is a set of data on magnetic fields of 770 mT, 420 mT, 200 mT, and 0 mT. In addition, in this work, the model capability to predict outside the training data range was evaluated. The selected datasets for prediction are 830 and 310 mT.

![Figure 1. SLFN of the ELM-based rheological model](image)

![Figure 2. Obtained flow curve on various magnetic field](image)
2.4. Simulation Setup

The simulation platform is Matlab® with a PC having a processor of 64 bit (3.4GHz). Then, the fitness or accuracy of the training and prediction performance will be assessed based on the root mean square error (RMSE), normalized RMSE (NRMSE) and coefficient of determination ($R^2$) with the equations of Equation (16), (17), and (18), respectively.

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i}^{N} (\tau_e - \tau_p)^2}
\]  

(16)

\[
NRMSE = \frac{RMSE}{(\tau_{max} - \tau_{min})} \times 100\%
\]  

(17)

\[
R^2 = 1 - \frac{\sum_{i}^{N} (\tau_e - \tau_p)^2}{\sum_{i}^{N} (\tau_e - \tau_{mean})^2}
\]  

(18)

where, $\tau_e$, $\tau_p$, $\tau_{max}$, and $\tau_{min}$ are measured, predicted, maximum measured, and minimum shear stress (kPa), respectively. The complete simulation setup is shown in Table 1. The logarithm normalization is carried out only for the shear rate.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hidden Nodes</td>
<td>20, 200, 500, 1 000, 2 500, 5 000, 7 000, 10 000</td>
</tr>
<tr>
<td>Activation Function</td>
<td>Hard Limit (HL), Sigmoid (Sig), and Sinusoid (Sin), Triangular basis function (Tribas), Radial Basis Function (Radbas)</td>
</tr>
<tr>
<td>Input weight and bias determination</td>
<td>Normal Distribution</td>
</tr>
<tr>
<td>Normalization method</td>
<td>Linear all, and log for shear rate only</td>
</tr>
</tbody>
</table>

3. RESULTS AND DISCUSSION

3.1. Effect on Accuracy at various activation functions

The comparison is carried out by comparing the experimental results with the produced shear rate at various activation functions. Table 2 describes the RMSE, NRMSE, and $R^2$ at linear (lin) and logarithm normalization (log). For the training case, the best accuracy is achieved by Hardlimit (ELM HL) for both log and lin case. Then, the second-best accuracy is Tribas (ELM Tribas) and followed by Radbas (ELM Radbas), Sinusoid (ELM Sin), and Sigmoid (ELM Sig). Both Lin and log has the same order in terms of accuracy. For linear case, this result is consistent with the result from the previous paper [25]. Meanwhile, the biggest gap between log and lin is achieved by ELM Sin. In other words, the ELM Sin gains the highest improvement comparing other activation functions. If the improvement is calculating using percentage as shown in Equation (19) where $RMSE_{Lin}$ is RMSE at Lin case and $RMSE_{Log}$ is RMSE at Log, the highest value achieved by Hardlimit followed by Tribas, Radbas, Sin, and Sig.

\[
\Delta E\% = \frac{(RMSE_{Lin} - RMSE_{Log})}{RMSE_{Lin}} \times 100\% 
\]  

(19)

<table>
<thead>
<tr>
<th>Schemes</th>
<th>Hidden Nodes</th>
<th>Log (kPa)</th>
<th>Lin (kPa)</th>
<th>Lin-log (kPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard Limit</td>
<td>10000</td>
<td>0.00</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>Sigmoid</td>
<td>10000</td>
<td>1.19</td>
<td>1.91</td>
<td>1.91</td>
</tr>
<tr>
<td>Sinusoid</td>
<td>10000</td>
<td>0.38</td>
<td>0.55</td>
<td>0.55</td>
</tr>
<tr>
<td>Tribas</td>
<td>10000</td>
<td>0.07</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Radbas</td>
<td>10000</td>
<td>0.32</td>
<td>0.45</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Hard limit and Tribas can achieve the best accuracy because of the simplicity and straightforwardness nature of the equation. Meanwhile, Sinusoid and Radbas have almost the same behavior, which is oscillation. The oscillation will be shown at linear normalization scheme and can be eliminated while using logarithm normalization. For example, Figure 3 shows the visual comparison between ELM Sin
Lin and ELM Sin Log. While the oscillation nature occurs at lin scheme, the log scheme can eliminate it and make the trend to follow the experimental results closely.

![Image](image1)

**Figure 3. The simulation results of ELM Sin for (a) Linear and (b) Logarithm normalization cases**

In terms of the testing accuracy, while predicting unknown data sets, the sigmoid has the best accuracy for log and lin test as shown in Table 3. ELM Sig is also can be considered as the most generalized or the least overfitting compared to others. The overfitting or not of the model can be checked in $|\text{RMSE}_p - \text{RMSE}_T|$ where $\text{RMSE}_p$ is RMSE for the prediction case and $\text{RMSE}_T$ is RMSE for training case. Also based on this criteria, the most overfitting model is ELM HL for log case and ELM Tribas for lin. Both models are the most overfitting for log and lin cases. As discussed before, the simple calculation nature may be able to produce high accuracy for training case. However, the simplicity can also cause the model cannot accurately predict the unseen data. This nature of ELM HL also occurs in the previous works [20],[25]. The highest improvement values after the application of the log scheme are achieved by ELM Tribas. Although it can be considered as the least overfitting models, $\Delta E\%$ has shown the best value which is about 77.10 %.

| Scheme     | Hidden Nodes | Log RMSE (kPa) | Log NRMSE (%) | Lin RMSE (kPa) | Lin NRMSE (%) | $|\text{RMSE}_p - \text{RMSE}_T|$ (kPa) | $|\text{RMSE}_p - \text{RMSE}_T|$ (%) | $\Delta E\%$ |
|------------|--------------|----------------|---------------|----------------|---------------|----------------------------------------|---------------------------------|------------|
| Hard Limit | 10000        | 2.32           | 3.31          | 2.32           | 2.21          | 1.26                                   | 1.43                            | 38.19      |
| Sigmoid    | 10000        | 1.13           | 1.61          | 0.07           | 1.43          | 2.04                                   | 3.04                            | 63.71      |
| Sinusoid   | 10000        | 1.66           | 2.35          | 1.26           | 1.86          | 2.65                                   | 3.04                            | 54.36      |
| Tribas     | 10000        | 1.66           | 2.38          | 1.59           | 2.53          | 3.62                                   | 2.21                            | 77.10      |
| Radbas     | 10000        | 1.70           | 2.43          | 1.38           | 1.81          | 2.58                                   | 1.06                            | 48.73      |

### 3.2. The effect on the computational cost

The computational cost can be analyzed from the point of view the need for the hidden node number to achieve a certain accuracy. The higher the hidden node number means that more resource is needed to reserve the memory to calculate $H^\dagger$ using SVD. In other words, the hidden node number represents the computational cost because the higher the hidden node number, the higher the time needed for the training process. The detailed of the efficiency of each model for the training process can be observed in Figure 4. For ELM HL, the ELM HL Log can reach less RMSE at less hidden node number compared to ELM HL Lin. The error also can be reduced to almost zero at 10000 hidden node number. The same phenomena also occur for the ELM Sin and Sig. The log scheme can reduce the requirement of the hidden node number to achieve a certain accuracy or RMSE.
Comparing the linear and logarithm normalized extreme learning machine in flow .... (Irfan Bahiuddin)

4. CONCLUSION

A model development is carried out to predict the shear stress of MR fluid as a function of shear rate and magnetic fields using ELM. In this work, a logarithm normalization method is investigated and compared with the linear normalization method. The evaluations are carried out at different activation function settings, which are Hard limit, sinusoid, sigmoid, triangular basis, and radial basis functions. The effect of the application logarithmic normalization can improve all the ELMs accuracy, especially ELM Sin which decreases the RMSE about 4 kPa. Meanwhile, in terms of prediction, the log normalization scheme can reduce the accuracy with the highest at RMSE of 1.9 kPa for ELM Tribas. The logarithm normalization can also reduce the computational burden by reducing the hidden node number to achieve a certain accuracy. The normalization methods can still be further investigated by applying the normalization method to either both or one of the inputs, which are magnetic field and shear rate and output. The evaluation for each region range can also be carried out to access the model capability for each low and high region of shear rate and magnetic field. Some scenario of the testing cases also needs to be carried out to check the model capability to predict unlearn datasets, such as interpolation case and extrapolation case. The same method can also be applied to other samples to check the consistency of the effect of the activation and normalization methods.

ACKNOWLEDGMENT

Authors acknowledge Universiti Teknologi Malaysia for the funding support through GUP Grant (Vote No: 15J44) and PDRU Grant (Vote No: 04E02).

REFERENCES