Comparing Neural Network Approaches in Modelling a Chromite Wet High-Intensity Magnetic Separator

J. Pretorius, A.F van der Merwe, F.H. Conradie and J. Cronje

***Abstract*—** Wet High Intensity Magnetic Separators (WHIMS) are increasingly used for the beneficiation of low-grade chromite ore, as chromium supply has been placed under significant pressure in recent years. The operational parameters of a WHIMS significantly affect its selectivity non-linearly. Statistical and mechanistic models have thus far only accounted for a limited number of variables, therefore, disregarding the majority of interactions between operational parameters and the combination thereof on the WHIMS's selectivity. In this paper, neural network approaches are compared for different WHIMS configurations, such that the versatility and validity of a WHIMS soft-probe control strategy can be determined. The different configurations are: single-stage and double-stage WHIMS. Furthermore, the training error is also varied for each configuration. Three back-propagation algorithms are investigated, namely: Levenberg-Marquart (LM), Bayesian Regularization (BR) and the Scaled Conjugate Gradient (SCG) algorithms. The best performing neural network has the lowest relative mean square error (MSE), whilst at the same time exhibits the highest relative coefficient of multiple determination R^2 for the output variables. Training data is obtained from experiments performed on a pilot WHIMS. The results show that the best performing neural network model for a single-stage WHIMS consists of 5 hidden neurons, trained using the BR algorithm. Furthermore, the best performing neural network model for a double-stage WHIMS consists of 5 hidden neurons, trained using the LM algorithm. From the comparison of architectures and training algorithms, it is concluded that a neural network model is certainly a viable method of controlling a chromite WHIMS.

***Index Terms*—**Chromite ore, artificial neural networks, number of hidden neurons, Wet High-Intensity Magnetic Separators

I. Introduction

Chromium ore, also termed chromite, is the main source of chromium metals and chemicals. Southern Africa contains approximately 90% of the world's chromite, with the majority, an estimated 3.1 billion tons, concentrated in the Bushveld Igneous Complex (BIC) of South Africa [1]. The production of ferrochrome requires chromium as constituent, the majority of which is used to produce stainless steel. However, the global supply of chromium has recently come under significant pressure due to the high demand for ferrochrome and stainless steel in addition to the inefficient beneficiation process of chrome resulting in greater waste and less output [2].

The inefficiency of chrome beneficiation is caused by the production of fine particles during the ore's comminution. These particles may constitute up to an estimated 25% of the ore's original chromium content, and consequently a significant

portion of potential revenue is lost with an added side effect to the natural environment and human life as chromium is toxic in high concentrations [3]. The particles are discarded from the beneficiation process as tailings and get stored in tailing dams.

As a result, tailing recovery companies have developed plants to recover the discarded chromium. Conventional separation techniques become increasingly ineffective in concentrating the chromium as the particles' sizes decreases [4]. Nonetheless, the application of Wet High-Intensity Magnetic Separators (WHIMS) has been shown to increase the chrome-to-iron ratio and the chromium oxide concertation more effectively at smaller particle sizes compared to other techniques [5]–[7].

The operational variables of a WHIMS strongly affect its selectivity (performance) variables [8], therefore, the control and optimization of a WHIMS significantly impacts the plant's economy. Poor WHIMS performance is often caused by high feed flow rates that overloads the volumetric capacity of the separator [9]. However, the modelling and control of a WHIMS is complex, as operational and selectivity variables exhibit strong non-linearity [10].

Mechanistic and statistical models derived for the control of WHIMSs are complex and not necessarily applicable to the WHIMS as a control strategy [9]. Nonetheless, machine learning models have recently proven promising results when an artificial neural network was shown capable of predicting a single-stage WHIMS's product quality using its input and operating variables to a modest degree of accuracy [11].

Inspired by the biological neural network, artificial neural networks (ANN) are mechanisms capable of approximating universal functions with any given precision by learning the relationships between the variables of the data set [12]. Artificial neural networks are thus used as arbitrary function estimators [13]. To apply artificial neural networks to a WHIMS, the network is first trained (learning) with a representative data set (input and output data), whereafter any input data can be used to predict output data. Nonetheless, the accuracy of the neural network's predictions determines

Manuscript received October 25, 2019.
The work was supported and funded in part by Minprotech.
J. Pretorius is with the North-West University, Potchefstroom
A.F. van der Merwe is with the North-West University, Potchefstroom
F.H. Conradie is with the North-West University, Potchefstroom
J. Cronje is with Minprotech, Rustenburg

whether it can be successfully applied as a soft-probe control strategy for a WHIMS. To obtain the best performing neural network, sufficient data, adequate network architecture and computational capacity are required [14].

Network architecture refers primarily to the number of hidden neurons. The number of hidden neurons is proportional to the complexity of the system. The network is either overtrained or undertrained if there are too many or too few hidden neurons [15]. Therefore, the exhaustive search technique is applied to obtain the best performing neural network's by varying the network's topology and training algorithm. The best performing network has the minimum average error, over the set of repetitions, for each number of the hidden neurons' iteration. Therefore, a global minimum is found by characterizing the data set to find its prevalent characteristics.

The exhaustive search method is not conventionally used as an enhancement technique, as it inefficiently determines the best network by operating too long to train and iterative through the set ranges. However, for smaller data sets as used in this paper, the training executes significantly faster relative to other training techniques, such as the genetic algorithm. Therefore, the best topology and algorithm can be determined for the pilot WHIMS (few but significant data points) and applied to the industrial WHIMS plant (many data points).

The aim of this study is to compare neural network approaches to different WHIMS configurations, such that the versatility and validity of a WHIMS soft-probe control strategy can be determined. The different configurations are single-stage and two-stage WHIMS. Furthermore, the training error is also varied for each configuration. Three back-propagation algorithms are investigated, namely the Levenberg-Marquart (LM), Bayesian Regularization (BR) and the Scaled Conjugate Gradient (SCG) algorithms. The best performing neural network has the lowest relative mean square error (MSE), whilst, the highest relative coefficient of multiple determination (R^2) for the output variables.

II. Experimental Procedure

A. Materials & equipment

The materials and equipment used to perform the experiments are listed in Table II.

A pilot-scaled WHIMS is used in all the experiments. The WHIMS's model number is RW9301102 and was built in and imported from Australia

The chromite ore is obtained from two mines, namely Mooinooi and Millsill. Ten sieves are used for the PSD (particle size distribution), ranging from 38 µm to 850 µm. A wet-basis PSD is performed for all the experiments.

TABLE I: Materials & Equipment

Equipment	Materials
Pilot-scale WHIMS	Chrome ore
10-way splitter	Filter paper
2-way splitter	Wash water
Drying oven	
Rubber hammer	
Filter press	
Sieves & shaker	
Mass scales	
Hand scoops	
Cylindrical containers (20 L)	
Cylindrical storage bin (100L)	

B. Method

The experimental procedure was performed twice. Firstly, the ore supplied by the mines is used as input to the WHIMS, resulting in product and tailing outputs (single stage). Secondly, the single stage's product is used as feed to the WHIMS, resulting in a final product and tailings outputs (two-stage).

The WHIMS's input parameters are feed density, feed flow rate, wash water flow rate and magnetic flux. The boundaries for each parameter of the pilot WHIMS is obtained and the test values chosen to include the upper and lower boundary. Furthermore, the remaining test values are selected such that a normal distribution for each input parameter is obtained. Hence, aiding in the prediction capabilities of the model.

The ore is received in slurry form and is thus first dried. Samples are taken from the dried ore to determine the d_{50} (particle size) and EPM (particle spread) from a PSD. The chrome to iron ratio and chromium oxide content was assessed using XRF (X-ray fluorescence) analysis. The ore is fed to the WHIMS where the feed density, feed flow rate, feed water flow rate and magnetic flux is varied individually from a basis. The product obtained after operating the WHIMS, is dried and samples taken again for PSD and XRF analysis. From the results, the recovery, grade and yield are calculated according to the Equations 1 to 3.

$$\text{Recovery} = \frac{\text{Mass of chrome in product}}{\text{Mass of chrome in feed}} \tag{1}$$

$$\text{Yield} = \frac{\text{Mass of product}}{\text{Mass of feed}} \tag{2}$$

$$\text{Product Grade} = \frac{\text{Mass of chrome in product}}{\text{Mass of product}} \tag{3}$$

III. Model Development

Both single- and double-stage experimental data is available for modelling the WHIMS. The single stage data, obtained from recent literature [11] is comparable to the double stage's data because both data sets were experimentally acquired from the exact same pilot WHIMS, experimental procedure and sample locations. The double stage's data comprises only this study's experimental results. Due to the comparable nature of the data, a third data set is introduced by combining the single- and double-stage data.

The three modelling data sets each comprise the same input

and output variables. The inputs are: d_{50} (particle size), EPM (particle spread), feed flow rate, feed density, wash water flow rate and magnetic flux. The outputs are: recovery, product grade and yield.

The exhaustive search technique is used to determine the training algorithm and number of hidden neurons configuration to develop the best performing neural network. The method entails iteratively training neural networks and varying the number of hidden neurons and training algorithm per iteration to obtain the best performing network. The number of hidden neurons investigated ranges between 1 and 25. Backpropagation training algorithms are used, specifically the LM, BR and SCG algorithms.

For each iteration, random initial weights and biases are assigned to the neural network, furthermore, the data set is randomly divided into three subsets. The subsets are used to perform different functions during the training process. Therefore, the network is trained repeatedly for each number of hidden neurons and training algorithm iteration. Hence, the best global performance is attained from the remaining repetitions' local performance. The number of repetitions range from 1 to 1000.

The best performing neural network is determined by evaluating the MSE and R^2 performance parameters. The MSE collectively accounts for the three output variables, whilst, the R^2 values for each output is determined. As mentioned, the best performing network is characterized by a minimum MSE value and maxima R^2 values (for each output). The MSE is calculated using Equation 4. The variable N refers to the number of data points and i refers to the iteration of the summation function. The variables Y'_i and Y_i represents the predicted and target values respectively.

$$\mathrm{MSE} = \sum_{i=1}^{N} \frac{(Y_i' - Y_i)^2}{N} \tag{4}$$

The procedure for determining the best network topology and training algorithm comprises three steps. Firstly, the exhaustive search technique is applied to train multiple neural networks, each with different number of hidden neurons and training algorithms, repeatedly.

Secondly, the mean of the MSE and R^2 performance parameters is computed over the entire repetition set for each number of hidden neurons iteration. The best performing number of hidden neurons is selected. By using the mean performance parameters, the characteristics of the data set is incorporated more effectively, as the data set is relatively small.

Thirdly, the best performing repetition for the selected number of hidden neurons is identified and the corresponding neural network extracted from the model's object arrays, hence, determining the best performing neural network for the specific training algorithm and data set used.

Step 2 and Step 3 are performed for all three training algorithms; hence, best training algorithm and corresponding number of hidden neurons are determined to ultimately attain the best performing neural network. All three steps are repeated for the single-stage, double-stage and combined stages experimental WHIMS data sets.

The method is programmed in MATLAB, using the Statistics and Machine Learning Toolbox. The time taken to train the networks is relatively short, as there are not many data points. However, the training time will increase as more data is added. Nonetheless, training can be started with the previous model's weights to increase the rate of training with the new data and prevent local minima convergence.

IV. Results & Discussions

The best performing neural network results for the exhaustive search method are listed in Table II.

TABLE II: Best performing neural networks for each algorithm and WHIMS configuration

No. Stages	Train Algorithm	No. Hidden Neurons	MSE	R^2 Recovery	R^2 Grade	R^2 Yield
One	BR	5	1.10E-05	0.999	0.990	0.998
One	SCG	6	1.23E-04	0.971	0.975	0.979
One	LM	3	9.61E-05	0.975	0.969	0.990
Two	BR	3	9.15E-05	0.993	0.992	0.988
Two	SCG	10	1.32E-04	0.986	0.977	0.992
Two	LM	5	8.77E-05	0.989	0.992	0.994
Combine	BR	2	5.84E-04	0.976	0.950	0.981
Combine	SCG	10	2.36E-04	0.996	0.933	0.997
Combine	LM	6	1.85E-04	0.996	0.973	0.998

The best performing neural network model for the single-stage WHIMS constitutes 5 hidden neurons, trained using the BR algorithm. The neural network's MSE is 1.10E-05 and the R2 values for the recovery, grade and yield are 99.9%, 99.0% and 99.8% respectively. The MSE is the lowest and the outputs' R^2 values the highest relative to the best performing LM- and SCG-trained neural networks.

The relationship between the MSE and number of hidden neurons is show in Fig. 1. The blue lines represent the MSE and the red, yellow and green lines represents the R^2 value for recovery, product grade and yield respectively. The minimum MSE value is evaluated in the range before the spike at 7 hidden neurons, as the network is less likely to be overtrained.

The prediction (red dots) and target (blue circles) results are plotted for each data point in Fig. 2, 3 and 4 for recovery, product grade and yield, respectively.

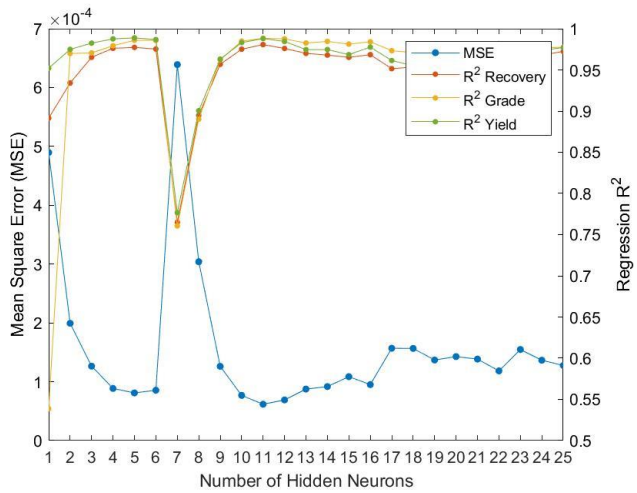

Fig. 1. MSE and regression values versus the number of hidden neurons for the BR algorithm

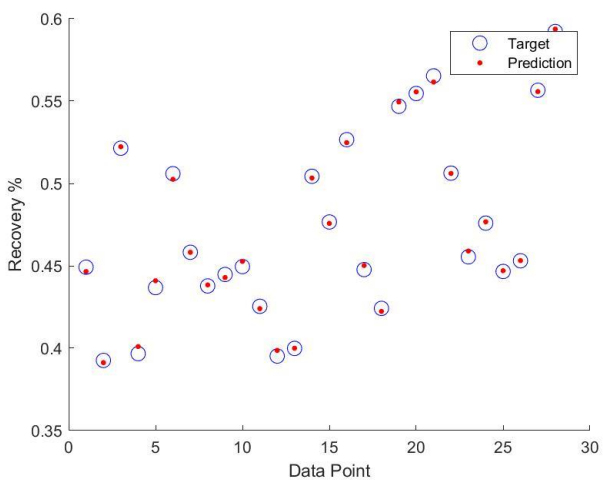

Fig. 2. Targets and predictions for the chrome recovery %

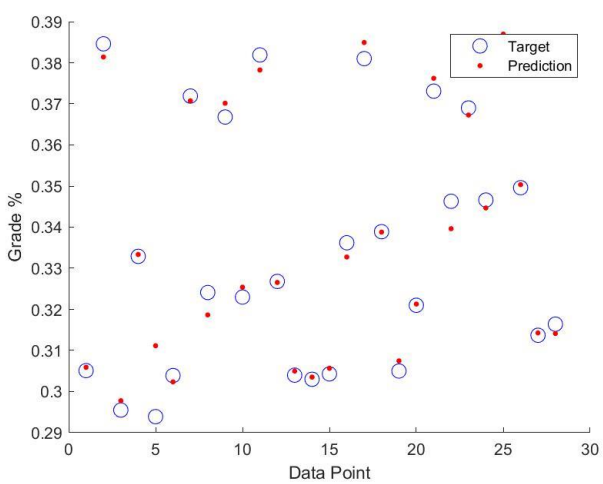

Fig. 3. Targets and predictions for the product grade %

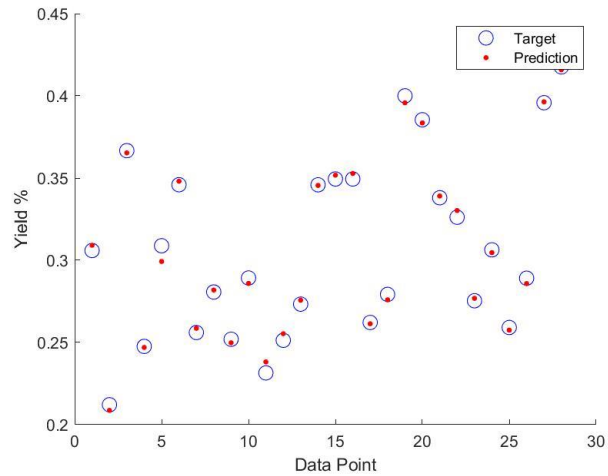

Fig. 4. Targets and predictions for the chrome yield %

For the two stage WHIMS configuration, the best performing neural network model constitutes 5 hidden neurons and is trained using the LM algorithm. The neural network's MSE is 8.77E-05 and the R^2 values for the recovery, grade and yield are 98.9%, 99.2% and 99.4% respectively. The MSE is the lowest and the outputs' R^2 values are the highest relative to the best performing BR- and SCG-trained neural networks.

The relationship between the MSE and number of hidden neurons is show by Fig. 5. The prediction results are given by Fig. 6, 7 and 8 for recovery, grade and yield, respectively. The predictions for recovery and yield are extraordinarily accurate, whilst, the prediction accuracy for grade is satisfactory – relative to the locally converged performances acquired by the other repetitively trained neural networks.

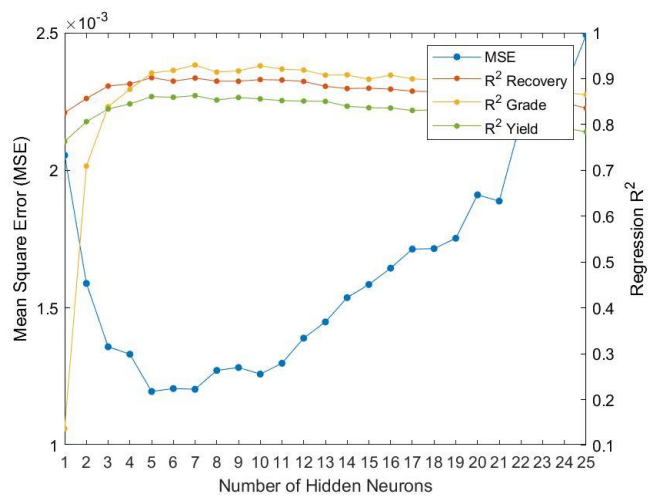

Fig. 5. MSE and regression values versus the number of hidden neurons for the LM algorithm

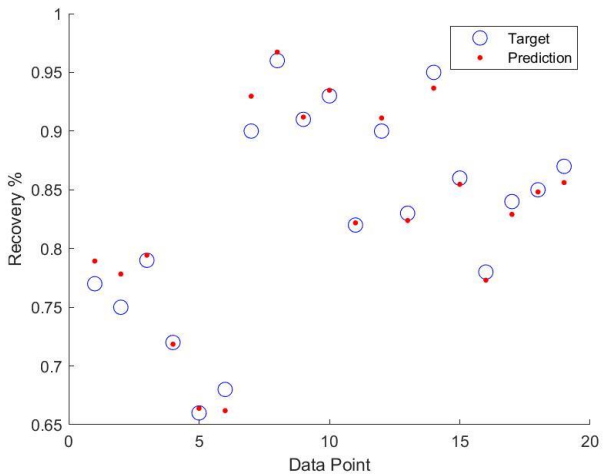

Fig. 6. Targets and predictions for the chrome recovery %

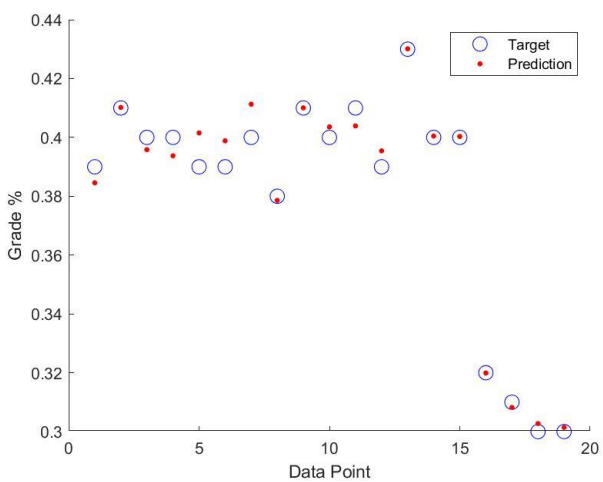

Fig. 7. Targets and predictions for the product grade %

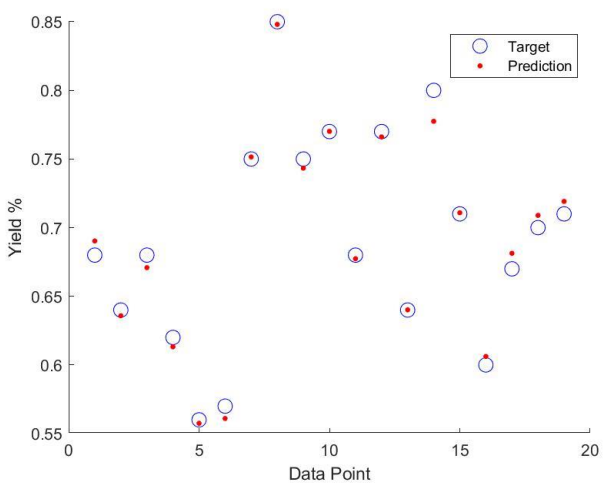

Fig. 8. Targets and predictions for the chrome yield %

The best performing neural network for the combined data set constitutes 7 hidden neurons, trained using the LM algorithm. The neural network's MSE is 1.85E-04 and the R^2 values for the recovery, grade and yield are 99.6%, 97.3% and 99.8% respectively. The MSE is the lowest and the R^2 values the highest relative to the best performing BR- and SCG-trained neural networks.

The relationship between the MSE and number of hidden neurons is show by Fig. 9. The prediction results are given by Fig. 10, 11 and 12 for recovery, grade and yield, respectively. The recovery and yield predictions demonstrate satisfactory performance. The product grade exhibits poor performance, as it has the lowest R^2 value and prediction accuracy in comparison to the recovery and yield outputs.

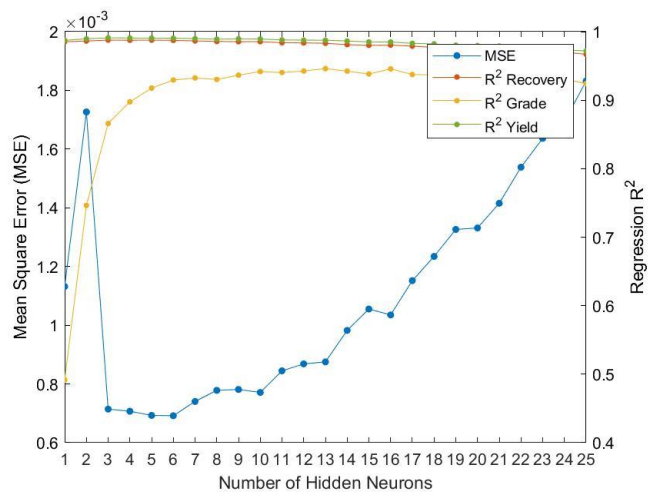

Fig. 9. MSE and regression values versus the number of hidden neurons for BR algorithm

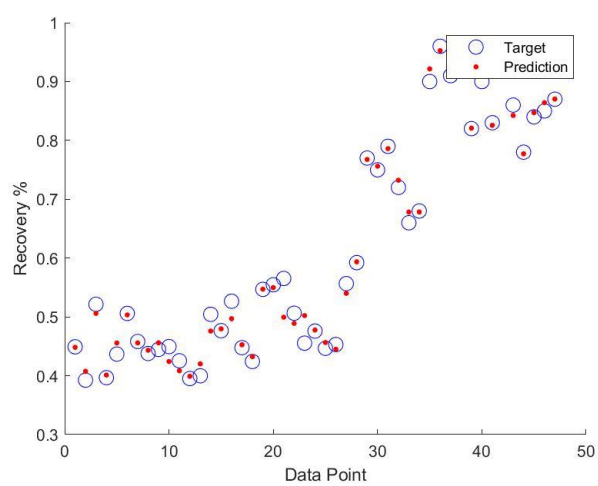

Fig. 10. Targets and predictions for the chrome recovery %

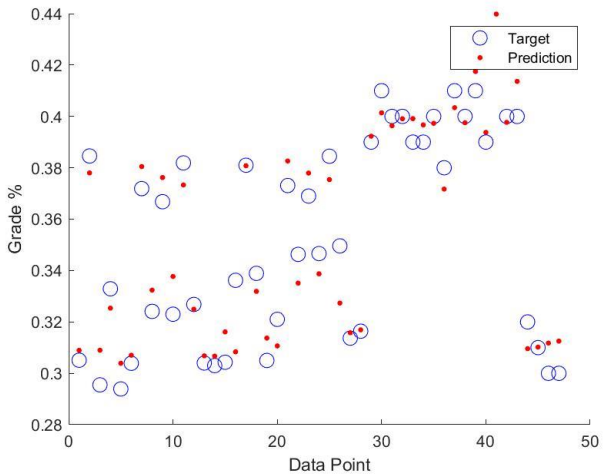

Fig. 11. Targets and predictions for the product grade %

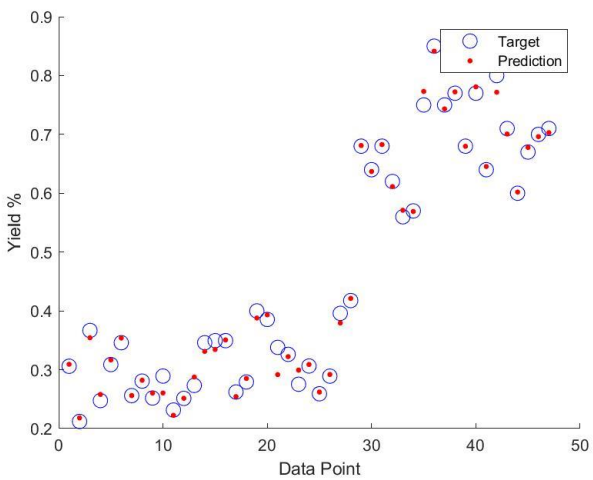

Fig. 12. Targets and predictions for the chrome yield %

V. Conclusions

The purpose of this paper was to compare different neural network approaches for the modelling of a WHIMS. The best performing training algorithms and architectures for a single-stage and two-stage WHIMS were determined. All training algorithms performed similarly well, with the exception of product grade predictions. The number of hidden neurons had a greater effect on the network's performance than the different algorithms. The LM algorithm was more versatile across the data sets compared to the SCG algorithm.

The results show that the best performing neural network model for a single-stage WHIMS consists of 5 hidden neurons, trained using the BR algorithm. Furthermore, the best performing neural network model for a double-stage WHIMS consists of 5 hidden neurons, trained using the LM algorithm.

The results agrees with the conclusion made by Reichel [11], that a single-stage WHIMS can be modelled using a neural network. This paper thus extends the conclusion to a two-stage WHIMS. Nonetheless, more experimental and plant data is required to determine the applicability of the experimental model to an industrial WHIMS, subject to internal and external disturbances.

From the comparison of network architectures and training algorithms, it is concluded that a neural network model is a viable method for the operational control of an industrial WHIMS. The performance of a WHIMS can be improved by implementing the appropriate control model, such as a soft-probe neural network model that predicts the output variables in real-time. Therefore, the model's application may allow improved recovery of chrome, as more small-sized chromite particles will be concentrated and not discarded as tailings to sludge dams. Hence, potentially aiding the alleviation of chrome's high demand in the long term and reducing the negative impact of chrome tailing particles on the environment.

References

[1] Y. R. Murthy, S. K. Tripathy, and C. R. Kumar, "Chrome ore beneficiation challenges & opportunities - A review," *Miner. Eng.*, vol. 24, no. 5, pp. 375–380, 2011. https://doi.org/10.1016/j.mineng.2010.12.001

[2] N. Aslan and H. Kaya, "Beneficiation of chromite concentration waste by multi-gravity separator and high-intensity induced-roll magnetic separator," *Arab. J. Sci. Eng.*, vol. 34, no. 2 B, pp. 285–297, 2009.

[3] S. Ozkan and B. Ipekoglu, "Investigation of environmental impacts of tailings dams," *Environ. Manag. Heal.*, vol. 13, no. 3, pp. 242–248, Aug. 2002. https://doi.org/10.1108/09566160210431042

[4] N. Gence, "Beneficiation of elazig-kefdag chromite by multi-gravity separator," *J. Eng. Environ. Sci.*, vol. 23, pp. 473–475, 1999.

[5] O. Abubakre, R. Muriana, and P. N. Nwokike, "Characterization and Beneficiation of Anka Chromite Ore Using Magnetic Separation Process," *J. Miner. Mater. Charact. Eng.*, vol. 06, no. 02, pp. 143–150, 2015. https://doi.org/10.4236/jmmce.2007.62012

[6] R. Rasool and H. Lieberwirth, "A continuum based numerical modelling approach for the simulation of WHIMS," *Miner. Eng.*, vol. 118, pp. 97–105, 2018. https://doi.org/10.1016/j.mineng.2017.12.020

[7] R. B. Rao, P. S. R. Reddy, S. Prakash, and M. I. Ansari, "Recovery of chromite values from chromite ore beneficiation plant tailings," *Trans. IIM*, vol. 40, no. 3, pp. 203–206, 1987.

[8] I. J. Corrans, W. A. Gilbert, K. S. Liddell, and R. C. Dunne, "The performance of an industrial wet high-intensity magnetic separator for the recovery of gold and uranium," *J. South African Inst. Min. Metall.*, vol. 84, no. 3, pp. 57–63, 1984.

[9] J. G. Rayner and T. J. Napier-Munn, "A mathematical model of recovery of dense medium magnetics in the wet drum magnetic separator," *Int. J. Miner. Process.*, vol. 69, no. 1–4, pp. 157–173, 2003. https://doi.org/10.1016/S0301-7516(02)00122-9

[10] W. W. Guo, M. Li, Z. Li, and G. Whymark, "Approximating nonlinear relations between susceptibility and magnetic contents in rocks using neural networks," *Tsinghua Sci. Technol.*, vol. 15, no. 3, pp. 281–287, 2010. https://doi.org/10.1016/S1007-0214(10)70062-6

[11] C. R. M. Reichel, A. F. Van Der Merwe, and J. Cronje, "A Predictive Model for Wet High Intensity Magnetic Separator (WHIMS) using Artificial Neural Networks," in *Eminent Association of Pioneers*, 2018.

[12] R. G. Kamp and H. H. G. Savenije, "Optimising training data for ANNs with genetic algorithms," *Hydrol. Earth Syst. Sci.*, vol. 10, no. 4, pp. 603–608, 2006. https://doi.org/10.5194/hess-10-603-2006

[13] M. Bahrami, M. Akbari, S. A. Bagherzadeh, A. Karimipour, M. Afrand, and M. Goodarzi, "Develop 24 dissimilar ANNs by suitable architectures & training algorithms via sensitivity analysis to better statistical presentation: Measure MSEs between targets & ANN for Fe–CuO/Eg–Water nanofluid," *Phys. A Stat. Mech. its Appl.*, vol. 519, pp. 159–168, 2019. https://doi.org/10.1016/j.physa.2018.12.031

[14] W. S. Sarle, "Neural Networks and Statistical Learning," in *SAS Users Group International Conference*, 1994.

[15] K. G. Sheela and S. N. Deepa, "Review on methods to fix number of hidden neurons in neural networks," *Math. Probl. Eng.*, vol. 2013, 2013 https://doi.org/10.1155/2013/425740