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.