Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. **00(0)** S1-S2 (2025)
Supplementary material

SUPPLEMENTARY MATERIAL TO

Optimizing ethylene plant utilities via Hybrid ANN and first-principles modelling

ALEKSA MILADINOVIĆ[1]*, MIRJANA KIJEVČANIN[2], JOVAN JOVANOVIĆ[2], SABLA ALNOURI[3], VLADIMIR STIJEPOVIĆ[1], MIRKO STIJEPOVIĆ[2]

[1]*Institute of Chemistry, Technology and Metallurgy, National Institute of the Republic of Serbia, University of Belgrade, Njegoševa 12, 11000 Belgrade, Republic of Serbia,* [2]*Faculty of Technology and Metallurgy, University of Belgrade, Karnegijeva 4, 11120 Belgrade, Republic of Serbia, and* [3]*Gas Processing Center, College of Engineering, Qatar University, Doha, Qatar.*

PROBLEM STATEMENT AND MODEL FORMULATION

Artificial neural network (ANN) is used to capture the system behavior (illustrated in Fig. 1 in the main body of the paper).

This configuration defines the input layer of the ANN with nine nodes, each corresponding to one of the input parameters. Specifically, the input data was normalized to the range [-1,1] using min-max normalization and fed into a feedforward neural network consisting of two hidden layers with twelve nodes each and an output layer with five nodes. The forward pass of the hidden layers of the network is calculated by applying a linear transformation followed by a non-linear activation function. The linear transformation consists of a weighted sum of the inputs plus a bias term:

$$Z = \left(\sum_{i=0}^{n} W_i \cdot X_i\right) + c \qquad \text{(S1)}$$

where Z is a pre-activation value, W is a weight matrix associated with the i neuron, c is a bias vector, and X is an input vector for the neuron i.

The used non-linear activation function (A) is the hyperbolic tangent function, which is implemented in the following equivalent form:

$$A = \frac{2}{1+e^{-2Z}} - 1 \qquad \text{(S2)}$$

The output layer consists of five nodes representing the required power of the three turbines ($W_{RT1,\mathrm{req}}$, $W_{RT2,\mathrm{req}}$ and $W_{RT3,\mathrm{req}}$) the temperature of the MP steam before mixing with the BFW and the mass flow rate of the BFW in the mixing process. This level only uses the linear transformation.

The neural network was trained with a data set that was split 80 % for training, 10 % for validation and 10 % for testing. The Levenberg–Marquardt

* Corresponding author. E-mail: aleksa.miladinovic@ihtm.bg.ac.rs

backpropagation algorithm, which is characterized by its efficiency and fast convergence in non-linear regression problems, was used for training. The number of training epochs was set to 10000 to ensure sufficient training iterations. The maximum absolute percentage error (MAPE) was used to evaluate the prediction accuracy, as it is a more representative measure of the model's performance, especially when the data contain variables with different magnitudes.

The performance of the neural network component during the training phase is illustrated in Fig. 1, demonstrating its effectiveness in learning the underlying patterns and relationships within the operational data.

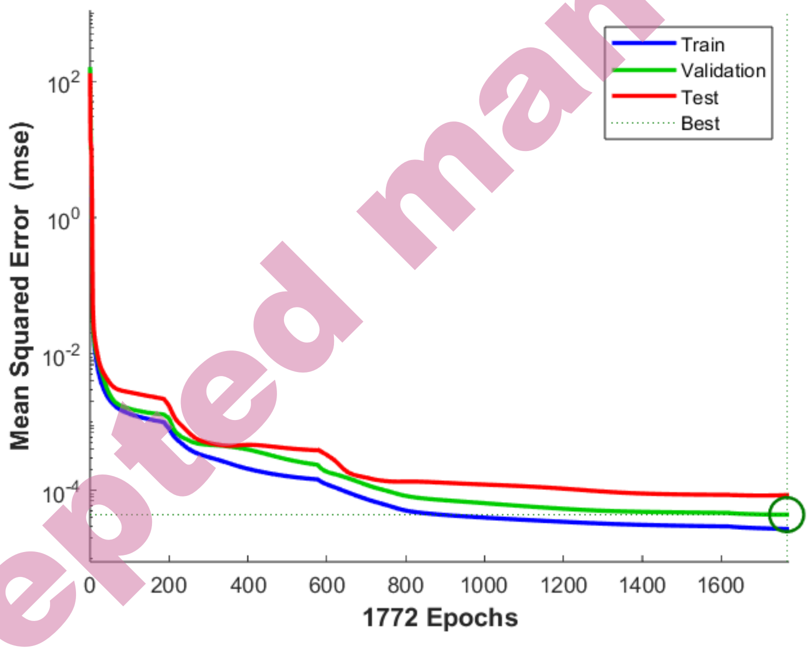

Fig. 1. Correlation plots of neural network predictions during training, validation, and testing.

The model was trained for 1772 epochs, monitoring the mean squared error (MSE) for the training, validation and test datasets. The training curve shows a rapid initial decrease in error, followed by a slower convergence phase, finally reaching a minimum validation error of 4.4136×10^{-5} at epoch 1766. The close agreement between the training, validation and test curves indicates good generalization and shows that the model was not over-fitted to the training data.

To further evaluate the prediction accuracy, the maximum absolute percentage error (MAPE) was calculated for each of the five output variables. The resulting MAPE values were 0.148 %, 0.221 %, 0.491 %, 0.023 % and 0.199 %, confirming the high precision of the model for all outputs.