Dear Reviewer,

We appreciate your comments and consider that the suggested modifications have contributed to improving the quality of our manuscript.

Answers to your comments are below.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

The manuscript describes application of neural network modelling approach to complex process such as flotation. The paper is interesting and deserve to be published in Hemijska Industrija journal, but there are several remarks that should be addressed:

1. Figure with input and output variables connected via neural network nodes and layers, would be of benefit for better paper idea presentation

**Answer**: A graph of a multilayer perceptron was incorporated into the manuscript, along with an explanation of the input variables and the output or response variables. It was indicated that the modeling and simulation were developed using the Keras library in Python, the activation functions and the learning method used were indicated.

1. How many experimental data is used for training neural network?

**Answer**: The size of the data set is approximately 1800 samples, so the training set is given by approximately 1260 samples.

1. What is the type of activation function used: linear, step, or …?

**Answer**: ReLU - Rectified Lineal Unit, for the neurons of the hidden layers and sigmoid for the output layer. This was stated in the manuscript.

1. What type of numerical method or software is used to estimate coefficients of activation function? Is it in-house or available on market?

**Answer**: Backpropagation, short for "backward propagation of errors," is the algorithm for supervised learning of artificial neural networks using gradient descent that was used. Given an artificial neural network and an error function, the method calculates the gradient of the error function with respect to the neural network's weights. The modeling, training, and testing of the neural network was developed using the Keras library in Python.

1. How do you explain that goodness of fitting R2 has relatively low value, while Mean Absolute Deviation (MAD) and Mean Squared Error (MSE) have also low values?

**Answer**: Since the activation function of the output layer is the sigmoid function, it is normal for the residuals to be low (especially the squared differences), since the difference between two values in the interval (0,1) and the size is calculated of samples "n" is large. Finally, considering the dynamics of the modeled process, the determination coefficients of the adjusted neural network models are significant for the authors.