Application of Neural Networks to Assess the Resource Value of Oil-Contaminated Waste Storage Facilities

A A Afinogentov¹, Yu A Bagdasarova¹, M Yu Derevyanov², Yu E Pleshivtseva²

¹Institute of Oil and Gas Technologies, Samara State Technical University, Molodogvardeyskaya Str., 244, Samara, 443100, Russia
²Department of Heat and Power Engineering, Samara State Technical University, Molodogvardeyskaya Str., 244, Samara, 443100, Russia

E-mail: pondex@yandex.ru

Abstract. The article presents a methodology for evaluating the efficiency of oil industry waste recycling systems using multi-layer artificial neural networks. As an indicator of the efficiency of the recycling system, the indicator of the resource value of oil-contaminated waste (OCW) is used. For training neural networks, the data sets are formed using the resource value assessment algorithm based on the Data Envelopment Analysis (DEA) method of multi-factor evaluation of the efficiency of production systems. The development and training of neural networks are performed using the free software Neuroph Studio. A comparative analysis of the quality of the assessment of the OCW resource value depending on the size and number of layers in a multi-layer neural network is carried out. The obtained results demonstrate the prospects of the proposed approach. Recommendations for improving the accuracy of resource value assessment by an artificial neural network are given.

1. Introduction
The current task of managing waste management systems for oil and gas production within the framework of the concept of a circular economy is to assess production and consumption waste as a production resource to be returned to circulation [1-6]. The efficiency of the reuse of oil and gas industry waste within the circular economy can be determined based on the environmental, economic, social, and technical complex indicators [1]. The considered efficiency indicators include heterogeneous parameters, for example, in [1] environmental indicators depend on parameters that characterize carbon emissions [2] and pollutants into the air, water and soil [3]; energy indicators estimate energy production and consumption in the process of waste treatment [4,5]; non-energy indicators consider the material balance, economic estimates, and social factors [6]. Thus, the definition of a generalized assessment of the efficiency of the oil and gas production waste disposal system requires a solution of the multi-factor analysis problem for determining the best possible scenario for waste processing that provides the maximum integral effect for all the above indicators. The described approach is developed by the authors in [7-10].

The procedure for multi-factor evaluation of the efficiency of the oil-contaminated waste processing system, which includes OCW storage facilities and technological installations for their utilization, is implemented in [7-9] based on the Data Envelope Analysis (DEA) method developed for evaluating the efficiency of production systems [10-12]. The proposed procedure uses the following
qualitative integrated indicators of the efficiency of the OCW recycling system: resource value, resource potential, environmental safety, energy efficiency. In [10] a comparative analysis using the cost-benefit method [13-15] was carried out for the best scenarios for the OCW processing, determined using the described procedure.

The integrated approach to assessing the efficiency of OCW processing systems presented in [7-10] is based on DEA models and digital technologies.

A promising approach at the current stage of digital technologies’ development is the use of artificial intelligence (AI), in particular the artificial neural networks (ANN), to predict and evaluate the effectiveness of OCW processing systems. The stages of training and testing on special data samples, the creation of which has its own features and approaches, are of fundamental importance for the effective use of ANN [16-17].

In this paper, we consider an approach to assessing the resource value of OCW storage facilities using artificial neural networks that allows to evaluate the effectiveness of the OCW processing system in general. A significant novelty lies in the approach to preparing a data set for training a neural network. It is suggested to form a set of “model” estimates of the OCW’ resource value using the DEA method according to the methodology described in [9].

2. Neural networks for the analysis of the resource value of OCW storage facilities

The proposed approach to the development of artificial neural networks for evaluating the effectiveness of OCW processing systems is based on the methods of system analysis and multi-criteria parametric evaluation of complex-structured systems proposed in [7-10].

In this paper, we consider the problem of creating an artificial neural network for evaluating the efficiency of the OCW processing system based on the complex indicator – the resource value of OCW storage facilities. Resource value represents a quantitative assessment of the physical and chemical composition and properties of OCW which determines the degree of suitability of waste in the estimated storage for the secondary use as material resources in recycling technologies.

The DEA model of resource value estimation proposed in the paper [9] is shown in Figure 1.

![Figure 1. DEA model for assessing the resource value of OCW storage facilities [9]](image)

In DEA method, the object under consideration is called a DMU (Decision Making Unit). The main input parameters that negatively affect the resource value \( R_{n}^V \) of the OCW storage (DMU) are the follows: the mass content of water \( x_{n1} \), asphaltenes and resins \( x_{n2} \), mineral parts and mechanical impurities \( x_{n3} \) and sulfur \( x_{n4} \). The output parameters that positively affect the assessment of the resource value of the OCW storage (DMU) are the follows: the mass content of light petroleum products in the OCW \( y_{n1} \) and the ratio \( y_{n2} = \frac{Q_{n}^{P}}{Q_{n}^{V}} \) of the mass of petroleum products \( Q_{n}^{P} \) to the mass of other components \( Q_{n}^{V} \). The calculation of the final assessments of the resource value of the OCW is based on the solution of a special mathematical programming problem providing the maximum of the objective function, which is the ratio of the sum of the weighted output parameters to the sum of the weighted input parameters of the DMU.

Table 1 shows the result of the assessment of the OCW resource value obtained on a data sample containing 110 OCW storage facilities in the Samara region. The Samara region can be considered as

| OCW storage | DMU | x_{n1} | x_{n2} | x_{n3} | x_{n4} | y_{n1} | y_{n2} |
|-------------|-----|--------|--------|--------|--------|--------|--------|
|             |     |        |        |        |        |        |        |

Note: Table 1 is not provided in the image.
a model region with a high density of the main OCW sources such as oil production, transportation, and processing facilities.

Table 1. Results of resource value evaluation for OCW storage facilities based on DEA method [9].

| OCW storage facilities | Input parameters | Output parameters | Resource value assessment |
|------------------------|------------------|-------------------|--------------------------|
|                        | water            | asphaltenes and resins | mineral parts | sulfur | petroleum products | $Q_v^n / Q_v^n$ | $RV_v$ |
| Storage №1             | 27.700           | 15.002             | 17.947         | 1.159  | 38.19             | 0.618       | 0.240 |
| Storage №2             | 26.241           | 20.000             | 19.770         | 2.372  | 31.62             | 0.462       | 0.258 |
| Storage №58            | 26.02            | 9.45               | 9.38           | 2.23   | 52.92             | 26.02       | 0.579 |
| Storage №59            | 34.22            | 5.27               | 10.61          | 1.69   | 48.20             | 34.22       | 0.789 |
| Storage №60            | 24.687           | 12.580             | 6.484          | 1.215  | 55.03             | 1.224       | 0.880 |
| Storage №109           | 32.155           | 5.353              | 17.480         | 2.207  | 42.804            | 32.155      | 0.748 |
| Storage №110           | 15.862           | 8.670              | 4.632          | 1.697  | 69.139            | 15.862      | 2.240 |

The resource value assessment obtained using the super efficiency DEA model characterizes the relative (in the comparison group) value of the OCW storage facility as a material resource for secondary use. The obtained estimates are distributed over the interval $[0, \infty]$, while all storages with assessment greater than or equal to 1 have a high resource value.

Determining the resource value of an individual OCW storage by the DEA method requires the formation of a data sample for compared objects, while the component composition and mass concentrations of components are typical for OCW oil-producing regions of Russia.

The analysis of the component composition and variations in the mass concentrations of components in the presented data sample allows us to conclude that it can be used as a training and test data sets for a multi-layer neural network to analyze the resource value of OCW storage facilities.

To create a training data set, 90 rows were randomly selected from the data sample containing 110 rows shown in Table 1. The remaining 20 rows were used to form the test data set shown in Table 2.

The implementation of neural networks is carried out using the free software Neuroph Studio – Java framework, which is a user interface that includes visualization tools, a constructor, and a library of neural network elements, as well as a module for training and testing a neural network on the corresponding data sets [18].

The structure of the neural network with two internal layers of neurons (layer 1 – 7 neurons, layer 2 – 15 neurons), implemented in the Neuroph Studio package, is shown in Figure 2.

In this paper we consider three neural networks with the following parameters:

- Neural network 1: 6 input parameters, 1 output parameter, 2 internal layers (layer 1 – 7 neurons, layer 2 – 15 neurons).
- Neural network 2: 6 input parameters, 1 output parameter, 2 internal layers (layer 1 – 7 neurons, layer 2 – 30 neurons).
- Neural network 3: 6 input parameters, 1 output parameter, 4 internal layers (layer 1 – 7 neurons, layer 2 – 31 neurons, layer 3 – 34 neurons, layer 5 – 31 neurons).

Neural networks have the following input parameters: the mass content of asphaltenes and resins, the mass content of water, the mass content of sulfur, the mass content of the mineral part and mechanical impurities, the mass content of light petroleum products, the ratio of the mass of petroleum products to the mass of the other components of the OCW. The output parameter is the value of the resource value of the OCW.

The proposed artificial neural networks were trained using the backpropagation method [19].
Figure 2. The structure of an artificial neural network in the Neuroph Studio package.

The training accuracy (TA) is defined according to the expression (1) as the root-mean-square error of the difference between the resource values obtained by DEA method for the training data sample and the results obtained by the artificial neural network. The training accuracy values are shown in Table 3.

The results obtained and the analyses of the quality of the OCW resource value assessment using each of the three neural networks are presented below.

3. Comparative analysis of neural networks
In this section, a comparative analysis of the accuracy of the resource values assessments using three neural networks is carried out on a test data set.

For a comparative analysis of accuracy, the calculated value of the following root-mean-square error is used (MSE):

$$
MSE = \frac{1}{N} \sum_{i=1}^{N} (r_i - \hat{r}_i)^2.
$$

The values of the maximum absolute (MAXAE) and maximum relative (MAXRE) errors are defined as follows:

$$
MAXAE = \max_{i=1,N} |r_i - \hat{r}_i|,
$$

$$
MAXRE = \frac{\max_{i=1,N} |r_i - \hat{r}_i|}{\hat{r}_i}.
$$

The values of the absolute mean (MAE) and the average relative (MRE) to the errors are considered according to the expressions:

$$
MAE = \frac{1}{N} \sum_{i=1}^{N} |r_i - \hat{r}_i|,
$$

$$
MRE = \frac{1}{N} \sum_{i=1}^{N} \frac{|r_i - \hat{r}_i|}{\hat{r}_i}.
$$
In (1)-(5) $r_i$ – the assessment of the OCW resource value determined by the neural network, $\hat{r}_i$ – the model assessment of the OCW resource value obtained based on the DEA method [9]; $N$ is the number of rows in the test data set.

The values $r_i$ and $\hat{r}_i$ for the test data set ($N = 20$) are shown in Table 2. Comparative quality indicators for considered ANN are presented in Table 3.

**Table 2.** Model (obtained by the DEA method) and neural network-defined assessments of the OCW resource value for the test data set.

| OCW storage facilities | Neural network 1 | Neural network 2 | Neural network 3 |
|------------------------|------------------|------------------|------------------|
|                        | $r_i$            | $\hat{r}_i$     | $r_i$            | $\hat{r}_i$     | $r_i$            | $\hat{r}_i$     |
| Storage №10            | 0.291            | 0.300            | 0.295            | 0.300            | 0.301            | 0.300            |
| Storage №12            | 0.730            | 0.731            | 0.565            | 0.731            | 0.701            | 0.731            |
| Storage №37            | 0.268            | 0.292            | 0.247            | 0.292            | 0.302            | 0.292            |
| Storage №42            | 0.536            | 0.527            | 0.532            | 0.527            | 0.548            | 0.527            |
| Storage №53            | 0.439            | 0.490            | 0.538            | 0.490            | 0.533            | 0.490            |
| Storage №96            | 0.661            | 0.648            | 0.524            | 0.648            | 0.644            | 0.648            |
| Storage №104           | 0.998            | 1.000            | 0.999            | 1.000            | 0.995            | 1.000            |

**Table 3.** Quality indicators for comparative analysis of neural networks.

| Quality indicator | Neural network 1 | Neural network 2 | Neural network 3 |
|-------------------|------------------|------------------|------------------|
| **Training stage**|                  |                  |                  |
| TA                | 0.0002           | 0.0001           | 0.0001           |
| **Testing stage** |                  |                  |                  |
| MAE               | 0.056            | 0.067            | 0.029            |
| MRE, %            | 9.2              | 10.7             | 5.7              |
| MSE               | 0.00814          | 0.00859          | 0.00149          |
| MAXAE             | 0.279            | 0.222            | 0.105            |
| MAXRE, %          | 39.6             | 31.5             | 26.0             |

4. Results and discussions

As a result of the research, a set of multilayer artificial networks was obtained, the training and testing of which was carried out using training and test data sets obtained by the procedure based on the DEA method of multifactorial assessment of the OCW storage facilities resource value.

The analysis of the accuracy of estimating the OCW resource value using artificial neural networks demonstrates the potential prospects of the proposed approach as well as leads to several conclusions:

- the model assessments of the resource value of the NSO (obtained by DEA method as the solutions of special mathematical programming problems) are suitable for the formation of data sets for the purpose of training multilayer neural networks.
- increasing the number of layers of a multi-layer artificial neural network allows one to improve the accuracy of the assessment: reducing the average absolute error MAE from 0.056 to 0.029, reducing the average relative error MRE from 9.2% to 5.7%, reducing the maximum absolute error MAXAE from 0.279 to 0.105, reducing the maximum relative error MAXRE from 39.6% to 26.0%.
- using a data sample containing 90 elements allows one to train a multi-layer neural network containing 100 artificial neurons.
The approach proposed in this paper to the creation of a data set for training an artificial neural network using model assessments of the resource value of OCW storage facilities, which were obtained based on the DEA method, is effective. The results obtained obey the well-known patterns of learning neural networks: an increase in the number of layers and the number of artificial neurons leads to an increase in the accuracy of the artificial neural network.

As recommendations for further research, we can offer an increase in the number of elements of the training data sample to improve the accuracy of the artificial neural network, as well as the implementation of the proposed approach to assess the effectiveness of the waste disposal system using such indicators as resource potential, environmental safety, and energy efficiency.

5. References

[1] Eleni Iacovidou, Costas A. Velis, Phil Purnell, Oliver Zwirner, Andrew Brown, John Hahladakis, Joel Millward-Hopkins and Paul T Williams 2017 Metrics for optimising the multi-dimensional value of resources recovered from waste in a circular economy: A critical review, J. Clean. Prod. vol 166 pp 910-938

[2] Christensen T H, Gentil E, Boldrin A, Larsen A W, Weidema B P and Hauschild M 2009 C balance, carbon dioxide emissions and global warming potentials in LCA-modelling of waste management systems Waste Manag. Res. vol 27 pp 707-715

[3] Allegreni E, Vadenbo C, Boldrin A, Astrup T F 2015 Life cycle assessment of resource recovery from municipal solid waste incineration bottom ash, J. Environ. Manag. vol 151 pp 132-143

[4] Arvidsson R, Fransson K, Froling M, Svans trom M and Molander S 2012 Energy use indicators in energy and life cycle assessments of biofuels: review and recommendations J. Clean. Prod. vol 31 pp 54-61

[5] Rita Puig, Pere Fullana-i-Palmer, Grau Baquero, Jordi-Roger Riba and Alba Bala 2013 Cumulative Energy Demand indicator (CED), life cycle based, for industrial waste management decision making Waste Manag vol 33, Issue 12 pp 2789-2797

[6] Fang K, Heijungs R, De Snoo G R 2015 Understanding the complementary linkages between environmental footprints and planetary boundaries in a footprint–boundary environmental sustainability assessment framework Ecol. Econ. vol 114 pp 218-226

[7] Derevyanov M Y, Pleshivtseva Y E and Afinogentov A A 2020 Simulation and Evaluation of the Efficiency of Oilcontaminated Wastes Recycling System IOP Conf.Ser.: Earth and Environ. Sci. 459 042058

[8] Derevyanov Maksim, Pleshivtseva Yuliya, Afinogentov Alexander, Mandra A and Pimenov Andrey 2019 Simulation and Multi-Objective Evaluation of Reuse Potential of Waste Recycling System for Oil And Gas Industry Proc. 21st International Conf. Complex Systems: Control and Modeling Problems CSCMP 2019 pp 429-434

[9] Pleshivtseva Yu E, Derevyanov M Yu, Kashirskikh D V, Pimenov A A, Kerov A V and Tyan V K 2018 Comparative evaluation of the reuse value of storage for oil-contaminated waste based on DEA method Neftyanye khozyaystvo - Oil Industry vol 11 pp 139-144

[10] Afinogentov A A, Derevyanov M Yu, Pleshivtseva Yu E 2021 Evaluation of the Technical and Economic Efficiency of Oil-Contaminated Wastes Recycling System Based on DEA-Method IOP Conf. Ser.: Earth Environ. Sci. 666 032032

[11] Chen Y and Du J 2015 Super-Efficiency in Data Envelopment Analysis (In: Zhu J. (eds) Data Envelopment Analysis. International Series in Operations Research & Management Science Springer, Boston, MA) vol 221

[12] Charnes A, Cooper W W and Rhodes E 1978 Measuring the efficiency of decision-making units”, European Journal of Operation Research vol 6 pp 429-444

[13] Begum R A, Siwar C, Pereira J J and Jaafar A H 2006 A benefitcost analysis on the economic feasibility of construction waste minimisation: the case of Malaysia Resour. Conserv. Recycl. vol 48 Issue 1 pp 86-98

[14] da Cruz N F, Simoes P and Marques R C 2014 Costs and benefits of packaging waste recycling
Acknowledgments
The reported study was funded by RFBR, project number 20-08-00353.