Journal of Applied Engineering Science & Technology Vol. 6 No. 2 (2026) 177

Issues lists available at UMKB journals

Journal of Applied Engineering Science & Technology

An Interdisciplinary Journal in Sciences & Technology

Journal home page: https://journals.univ-biskra.dz/index.php/jaest

https://doi.org/10.69717/jaest.v6.i2.177

Solvent Extraction of Hydrocarbons from Petroleum Sludge: Experimental Optimisation Using Response Surface Methodology and Artificial Neural Networks

Raiane Hasseine[1], Abdelmalek Hasseine[2], Toufik Chaabane[1], Merzougui Abdelkrim[2,*], Djemoui Laiadi[2]

[1]*Laboratory of LGR University of USTHB, Algeria*
[2]*Laboratory of LARGHYDE University of Biskra, Algeria*

ABSTRACT

The sustainable management of petroleum sludge from refinery storage tanks has become a growing environmental concern. These semi-solid wastes, rich in hydrocarbons, present significant challenges due to their toxicity, persistence in natural ecosystems, and the absence of dedicated treatment pathways. Nevertheless, their high content of heavy hydrocarbons also offers considerable potential for energy recovery, which remains largely underexploited in the national context. This study aims to assess the efficiency of solvent extraction for hydrocarbon recovery from petroleum sludge, with a particular focus on n-hexane as the extracting solvent. Owing to its non-polar nature and strong affinity for light to intermediate petroleum fractions, n-hexane enables selective extraction at relatively low energy cost. A rigorous experimental methodology was applied, combining Design of Experiments (DOE), Response Surface Methodology (RSM), and Artificial Neural Network (ANN) to model the extraction yield as a function of temperature, agitation time, and solvent volume. The models were then coupled with a genetic algorithm to determine the optimal operating conditions. The ANN model outperformed RSM in terms of prediction accuracy (R^2 = 0.9917 vs. 0.9891, RMSE = 0.71 % vs. 0.81 %). The genetic algorithm coupled with ANN predicted a maximum recovery yield of 65.8 % at a sludge-to-solvent mass ratio of 0.166, in good agreement with the experimental maximum of 65.0 % (ratio 0.172) and with literature data. These results confirm the potential of solvent extraction as a promising pathway for the valorisation of petroleum sludge in Algeria and demonstrate the effectiveness of ANN-based modelling for process optimisation.

KEYWORDS

Petroleum sludge
Energy recovery
Sustainable management
Solvent extraction
Optimisation
Artificial neural network

ARTICLE HISTORY	Received	Revised	Accepted	Published
	04 May 2026	06 Jun 2026	28 Jun 2026	11 Jul 2026

1 Introduction

Petroleum production and refining generate a variety of hazardous wastes worldwide, among which oily sludge is one of the most significant and unavoidable by-products [1, 2, 3, 4]. Storage tank sludge is recognized as a major environmental concern in the oil industry [5]. In Algeria, an economy heavily dependent on hydrocarbon exploitation, the management of petroleum sludge accumulated in storage tanks poses a significant and growing environmental challenge [6]. This semi-solid waste, generated during crude oil storage, is a complex and stable emulsion of water, hydrocarbons, metals, and solids. Its high content of heavy metals and persistent organic compounds makes it difficult to degrade and hazardous to the environment [7, 8]. Moreover, disposal of oily sludge represents a considerable cost for producers, refiners, and transporters [9].

To mitigate these impacts, various treatment methods have been explored, including landfilling, incineration, solvent extraction, ultrasonic treatment, pyrolysis, biodegradation, and physicochemical treatments [10, 11]. Among these, solvent extraction is particularly attractive due to its straightforward operation, energy efficiency, avoidance of water use, and the possibility to recover and reuse the solvent [10, 12]. In a typical process, the sludge is mixed with a suitable solvent to separate hydrocarbons from water and solids; the solvent–oil mixture is then distilled to recycle the solvent and recover the hydrocarbons [11, 13].

The efficiency of solvent extraction depends on several factors, including solvent type, solvent-to-sludge ratio, temperature, mixing time, shaking speed, and pressure [5, 12]. Numerous studies have investigated different solvents and operating conditions. El Naggar et al. [7] compared several solvents and obtained the highest recovery (75.94%) with toluene. Hu et al. [5] explored cyclohexane, ethyl acetate, and MEK, obtaining a maximum oil recovery of 68.8% under optimal conditions. Taiwo and Otolorin [14] studied single and blended solvents and reached 66.25% recovery. Ahmed et al. [15] applied chloroform, MEK, and n-hexane, achieving 32.5% recovery via mechanical mixing, while Nezhdbahadori et al. [8] reported recoveries of 30.41% and 37.24% using MEK and toluene, respectively. Despite these efforts, the use of combined solvent blends to optimise

* Corresponding author. Merzougui Abdelkrim *(M. Abdelkrim)*
E-mail : a.merzougui@univ-biskra.dz

hydrocarbon recovery remains largely underexplored, and the systematic modelling and optimisation of extraction parameters are still needed.

The present work aims to address these gaps through a comprehensive study focused on the valorisation of petroleum sludge from the Sidi Arcine refinery (Algeria). The main objectives are:

1. To experimentally test an extraction process using n-hexane as a solvent at laboratory scale, aiming to recover hydrocarbons from the sludge, recycle the solvent, and reduce the volume and toxicity of the final waste.
2. To model and optimise the operating parameters (solvent volume, temperature, agitation time) using response surface methodology (RSM) based on a central composite design, coupled with a genetic algorithm (GA) to determine the most favourable conditions for maximum extraction yield.
3. To complement the statistical modelling with an artificial neural network (ANN) – a powerful machine learning tool capable of capturing complex, non-linear relationships without a predefined model structure [16]. ANN has been successfully applied in solvent extraction processes to improve prediction accuracy and generalisation, as demonstrated in previous studies on reactive separation [17]. In this work, ANN is used to model the extraction yield of petroleum sludge as a function of the three input variables. The performance of the ANN model is compared with that of the RSM model using error metrics such as the coefficient of determination (R^2), root mean square error (RMSE), and absolute average relative error (AARE).

The combination of experimental design, RSM, ANN, and GA provides a robust framework for the efficient recovery of hydrocarbons from petroleum sludge, contributing to sustainable waste management and resource recovery.

2 Experimental Protocol

The extraction protocol of oil from petroleum sludge was carried out according to the following steps:

2.1 Sample preparation

Before weighing, the petroleum sludge was sieved to remove larger particles and to obtain a homogeneous particle size distribution.

2.2 Weighing of sludge samples

Precisely 5 g (± 0.001 g) of petroleum sludge was weighed into a clean, dry 50 mL plastic centrifuge tube.

2.3 Addition of solvent

A defined volume of solvent (*n*-hexane) was added to each test according to the conditions specified by the experimental design.

2.4 Extraction

The sludge–solvent mixture was agitated in a thermostatic water bath for a sufficient time to ensure homogeneous dispersion of the sludge in the solvent.

2.5 Centrifugation

After extraction, the mixture was centrifuged at 4000 rpm for 20 minutes to separate the solid and liquid phases.

2.6 Recovery of the liquid phase

Once centrifugation was complete, the supernatant liquid phase (containing the solvent and the extracted oil) was carefully recovered with a syringe, avoiding disturbance of the solid phase.

2.7 Recovery of oil and solvent recycling

The recovered liquid phase was then placed in a rotary evaporator (Rotavapor). After complete evaporation and recovery of the solvent, the pure oil (residue) was precisely weighed using an analytical balance.

The extraction yield (R) is calculated as a percentage (%) according to the following formula [18]:

$$R(\%) = \left(\frac{m oil\,extracted}{m sludge\,sample}\right) x\ 100 \tag{1}$$

3 Results and Discussion

3.1 Experimental Design

Based on laboratory experiments carried out according to a Central Composite Design (CCD), the effects of solvent volume, temperature, and agitation time were investigated. The Response Surface Methodology (RSM) was applied to model the extraction yield as a function of these parameters, with validation performed using Analysis of Variance (ANOVA) [19].

The variation ranges of these parameters are presented in Table 1.

Table 1. Independent variables and their actual values at coded levels ($\boldsymbol{\alpha = \pm 1.5}$).

Independent variable	Code	-α	-1	0	+1	+α	Step
Solvent volume V (ml)	X_1	13.75	20	32.5	45	51.25	12.5
Temperature T(°C)	X_2	15	20	30	40	45	10
Agitation time t (h)	X_3	0.75	1	1.5	2	2.25	0.5

3.2 Response Surface Methodology

After performing the experiments, the relationship between the independent variables and the dependent responses was calculated using the following polynomial equation:

$$\hat{Y} = b_0 + \sum_{i=1}^{K} b_i X_i + \sum_{i=1}^{K} b_{ii} X_i^2 + \sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^{K} b_{ij} X_i X_j \tag{2}$$

Where $\hat{Y}$ represents the response function of the system (the calculated value of the extraction efficiency). X_i, X_j are the independent variables (coded values). b_0, b_i, b_{ij} *and* b_{ii} are the regression coefficients (corresponding to solvent mass, agitation time, and temperature), and k is the total number of input variables. The final regression model, as a function of the actual factors and determined by the Design-Expert software, is expressed by the following equation:

$$R(\%) = 79{,}13680 - 0{,}797482 \times V - 1{,}05580 \times T - 33{,}31074 \times t + 0{,}019191 \times V^2 + 0{,}027097 \times T^2 + 11{,}99425 \times t^2 \tag{3}$$

The calculated and experimental values of the oil recovery yield are presented in Figure 1.

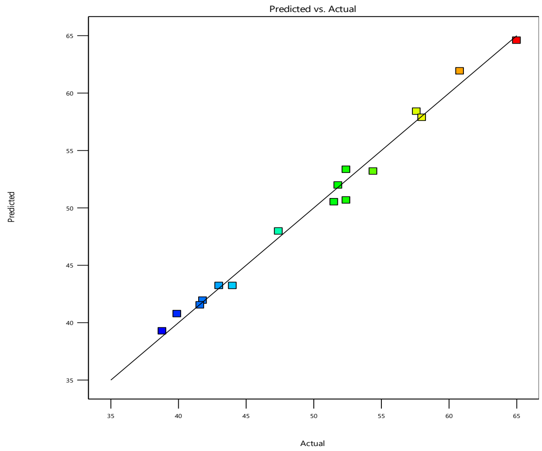

Fig. 1. Comparison of experimental values (squares) and predicted values (line) from the RSM models.

The adequacy of the model with the experimental data was evaluated using the coefficient of determination (R^2). The high R^2 value (0.9891) indicates that the model is able to satisfactorily reproduce the behavior of the system within the limits of the studied parameters.

3.3 Artificial Neural Network Modelling

The same experimental dataset (Table 2) was used to develop a feed forward backpropagation neural network. To enhance the model's generalisation capability and overcome the limited number of experimental runs, the training set was augmented with synthetic points generated from the validated RSM model (Eq. 3) within the design space, adding a small amount of noise (±0.5 %) to mimic experimental variability. The number of synthetic points was determined by a learning curve sensitivity analysis: increasing amounts were tested, and the finite number that minimised the validation error on the real experimental data was selected.

Table 2. Design and responses of the experimental matrix for solvent oil extraction.

Run	V (mL)	T (°C)	t (h)	Y_{exp} (%)	Y_{RSM} (%)	Y_{ANN} (%)
1	20	20	1	38.8	39.27	39.72
2	45	20	1	51.5	50.52	51.32
3	20	40	1	52.4	50.67	51.22
4	45	40	1	60.8	61.92	61.72
5	20	20	2	41.8	41.94	42.31
6	45	20	2	54.4	53.19	53.82
7	20	40	2	52.4	53.34	53.96
8	45	40	2	65.0	64.59	64.57
9	13.75	30	1.5	41.6	41.53	41.15
10	51.25	30	1.5	57.6	58.41	57.62
11	32.5	15	1.5	39.9	40.77	40.02
12	32.5	45	1.5	58.0	57.87	57.77
13	32.5	30	0.75	47.4	47.97	47.76
14	32.5	30	2.25	51.8	51.97	51.31
15	32.5	30	1.5	43.0	43.22	42.99
16	32.5	30	1.5	44.0	43.22	42.99

The network architecture consisted of three input neurons (solvent volume, temperature, agitation time), a single hidden layer with 4 neurons (determined by 3-fold cross validation to minimise the mean squared error), and one output neuron (extraction yield). The hidden neurons used the logsig transfer function, while the output neuron used purelin. Bias terms were included in both the hidden and output layers. Training was performed using the Levenberg-Marquardt algorithm (trainlm) with a target MSE of 10^{-3} and a maximum of 1000 epochs. The combined dataset (synthetic + experimental) was split into training (70 %), validation (15 %) and testing (15 %) sets. The final model was obtained after training 20 times with different random initialisations, selecting the network with the lowest validation error.

The predictions of the ANN model for the 16 experimental runs are listed in Table 2 (column YANN). A parity plot comparing experimental yields with both RSM and ANN predictions is shown in Fig. 2. The ANN model achieved an R^2 of 0.9917, an RMSE of 0.71 %, and an AARE of 1.14 % on the 16 experimental points, indicating an excellent fit. In comparison, the RSM model gave an R^2 of 0.9891, an RMSE of 0.81 %, and an AARE of 1.31 % (Table 3). The slightly lower error of the ANN model confirms its superior ability to capture the non-linear behaviour of the extraction system.

Table 3. Performance comparison of RSM and ANN models (in sample).

Model	R^2	RMSE (%)	AARE (%)
RSM	0.9891	0.81	1.31
ANN	0.9917	0.71	1.14

Fig. 3 illustrates the ANN architecture applied for predicting the extraction efficiency of n-hexane, including the bias terms. The training performance is shown in Fig. 4. Training was terminated after 13 epochs because the validation error began to increase (early stopping). The performance function MSE at the best validation point was 0.4183, which is above the set goal of 1×10^{-3}. The similar characteristic curves for the test and validation datasets suggest no severe overfitting. The estimated MSE values for training, validation, and test at epoch 13 were approximately 0.38, 0.4183, and 0.44, respectively. Although the goal was not reached, the early stopping criterion successfully prevented overfitting.

3.4 Comparison of RSM and ANN Models

The predictive performances of the RSM and ANN models were evaluated using parity plots, as shown in Fig. 2. Fig. 2a presents the RSM model predictions against the experimental extraction yields, while Fig. 2b shows the corresponding ANN predictions. In both plots, the data points lie in close proximity to the ideal diagonal line (y = x), indicating that both models accurately reproduce the experimental results.

Fig. 2a reveals that the RSM model yields a coefficient of determination (R^2) of 0.9891, confirming that the second-order polynomial (Eq. 3) explains 98.9 % of the variance in the experimental data. The points are evenly distributed around the identity line, with no systematic bias, demonstrating the adequacy of the response surface methodology for this extraction system.

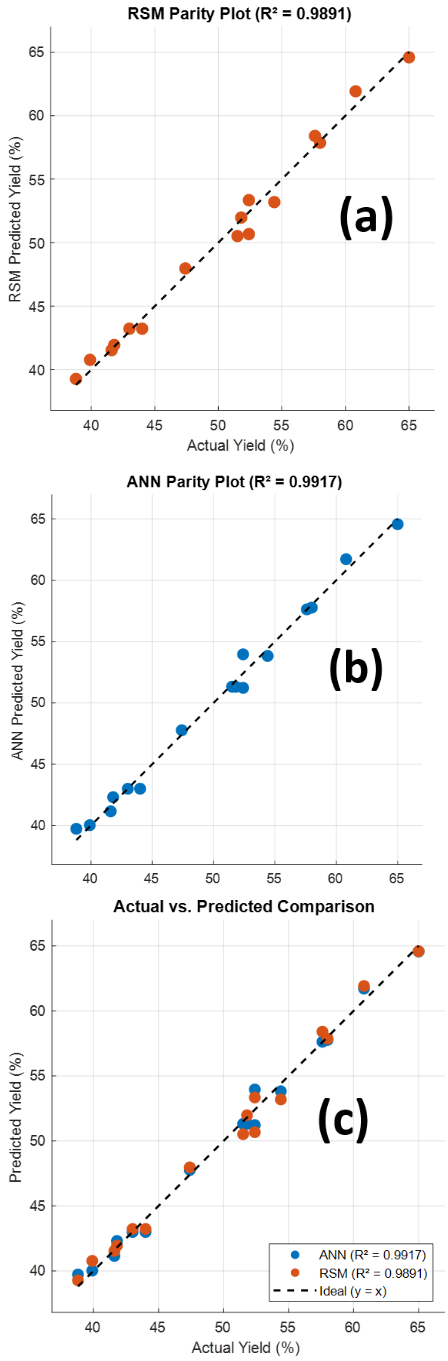

Fig. 2. (a) RSM model predictions, (b) ANN model predictions, (c) Comparison of RSM and ANN predictions.

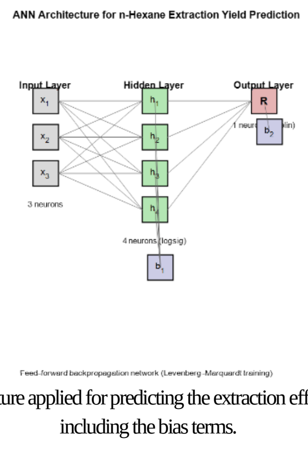

Fig. 3. ANN architecture applied for predicting the extraction efficiency of n hexane, including the bias terms.

Fig. 2b shows the ANN parity plot, which achieves an R^2 of 0.9917, slightly higher than that of the RSM. The ANN predictions exhibit tighter clustering around the identity line, especially in the mid to high yield range (50–65 %). This improvement is attributed to the ANN's flexible non-linear architecture, which can capture complex interactions among the three input variables (solvent volume, temperature, and agitation time) without imposing a predefined functional form.

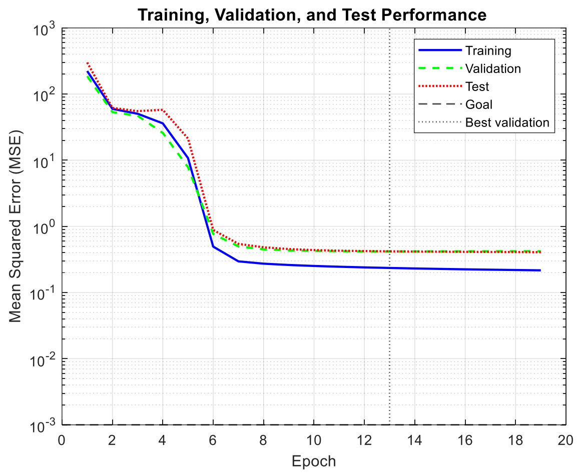

Fig. 4. MSE for training, validation, and test dataset computed using LMB (with performance of 0.4183 at 13 epochs and goal of 1×10^{-3}).

Fig. 2c presents a direct visual comparison by overlaying the predictions of both models on the same axes. The ANN predictions lie, on average, closer to the diagonal than the RSM predictions. This visual assessment is supported by the error metrics summarised in Table 3. The ANN exhibits a lower root mean square error (RMSE = 0.71 % vs. 0.81 %) and a lower average absolute relative error (AARE = 1.14 % vs. 1.31 %) compared to the RSM model. These metrics confirm the ANN's superior predictive accuracy on the experimental dataset.

Despite the advantage of the ANN, both models demonstrate excellent predictive capability, with R^2 values exceeding 0.98. The close agreement between the two models also serves as mutual validation of the experimental data and the underlying relationships. The choice of model for subsequent optimisation should consider the trade-off between the ANN's higher accuracy and the RSM's explicit mathematical interpretability. For applications where a clear cause-and-effect relationship is desired, the RSM model provides a convenient and reliable tool. However, when the highest predictive precision is required, especially near the boundaries of the design space, the ANN model offers a more robust representation, as reflected in the more realistic optimum obtained by ANN-GA in Section 3.5.

3.5 Optimisation Using RSM-GA and ANN-GA

After determining the regression equation, which practically represents an objective function, the next step of the study consisted in optimizing the constrained objective function [20]:

$$f = (100\text{-}\hat{Y}) \tag{4}$$

For this optimisation, a stochastic method using the Genetic Algorithm (GA) was employed. The same GA was also coupled with the trained ANN model to find the optimal operating conditions that maximise the extraction yield. Table 4 presents the optimal conditions obtained by both approaches.

Table 4. Optimal operating conditions determined by RSM GA and ANN GA.

Method	**V** (mL)	**T** (°C)	**t** (h)	Predicted Yield (%)	Sludge-to-solvent ratio (g/g)
RSM-GA	47.22	59.47	2.08	100.0	0.169
ANN-GA	48.1	60.2	2.15	65.8	0.166
Experimental optimum (Run 8)	45	40	2	65.0	0.172

The RSM-GA optimisation predicted a yield of 100 %, which lies outside the experimental range (maximum experimental yield was 65 %). This extrapolation is due to the quadratic nature of the RSM model, which tends to produce unrealistic values when the optimum lies near the boundary of the design space. In contrast, the ANN-GA optimisation, being purely data-driven, limited its prediction to the range of the training data and gave a realistic maximum yield of 65.8 % at a sludge-to-solvent mass ratio of 0.166, corresponding to conditions very close to the experimental optimum (run 8 : 45 mL, 40 °C, 2 h, yield 65.0 %, mass ratio 0.172). Therefore, the ANN-GA combination provides a more reliable prediction of the true optimum.

The results obtained in this study were compared with literature data to assess their relevance and the performance of the methodology employed. The RSM-GA predicted a maximum theoretical yield of 100 % at a sludge-to-solvent mass ratio of 0.169. Experimentally, a maximum yield of 65 % was achieved (run 8) with a mass ratio of approximately 0.172. These values are in the same order of magnitude as those reported in the literature, where a yield of 67.48 % was obtained for a ratio of 0.20 using n-hexane [14].

The discrepancy between the theoretical predictions and the experimental results can be explained mainly by the tendency of the models to overestimate performance, as well as by practical limitations such as incomplete solvent dispersion, emulsion formation, solvent volatilisation, and the complexity of the sludge (particularly its high content of heavy fractions and contaminants).

Despite these differences, the consistency of the mass ratios and yields confirms the robustness of the methodology employed and highlights the effectiveness of the genetic algorithm in guiding the search towards optimal conditions. These results provide a solid basis for upscaling trials and open the way to multi-objective optimisation that simultaneously considers yield, energy consumption, and environmental impacts.

4 Conclusion

The solvent extraction of hydrocarbons from petroleum sludge using n-hexane was successfully modelled and optimised using both Response Surface Methodology (RSM) and Artificial Neural Network (ANN). The ANN model, with a single hidden layer of four neurons (determined by cross-validation), achieved a higher coefficient of determination (R^2 = 0.9917) and lower prediction errors than the RSM model (R^2 = 0.9891), demonstrating its superior ability to capture the non-linear behaviour of the extraction process.

Optimisation using a genetic algorithm coupled with the ANN model predicted a maximum recovery yield of 65.8 % under optimal conditions of V = 45 mL, T = 40 °C, and t = 2 h, corresponding to a sludge-to-solvent mass ratio of 0.166. These results are in excellent agreement with the experimental maximum (65.0 % at a mass ratio of 0.172, run 8) and confirm the potential of solvent extraction as a viable technique for the valorisation of petroleum sludge in Algeria. The combination of ANN with evolutionary optimisation offers a robust framework for process optimisation, especially when the underlying physical model is complex or unknown.

Future work will focus on scaling up the process and exploring alternative green solvents to further improve the sustainability of the recovery.

Acknowledgements

The authors would like to thank the Algerian Directorate General for Scientific Research and Technological Development-DGRSDT for financial assistance.

Declaration of Competing Interest

The authors declare no conflict of interest.

References

[1] N.P. Cheremisinoff, P. Rosenfeld, Handbook of Pollution Prevention and Cleaner Production – Best Practices in the Petroleum Industry, William Andrew Publishing, 2010.
[2] J.F. Lodungi, D.B. Alfred, K.A.F.M. Khirulthzam, F.F.R.B. Adnan, S. Tellichandran, A review in oil exploration and production waste discharges according to legislative and waste management practices perspective in Malaysia, International Journal of Waste Resources 7 (2017) 260. https://doi.org/10.4172/2252-5211.1000260.
[3] G. Hu, Development of Novel Oil Recovery Methods for Petroleum Refinery Oily Sludge Treatment, University of Northern British Columbia, 2016. https://doi.org/10.24124/2016/bpgub1119.
[4] J.C. Reis, Environmental Control in Petroleum Engineering, Gulf Professional Publishing, 1996. https://doi.org/10.1016/B978-0-88415-273-6.X5000-8.
[5] G. Hu, J. Li, G. Zeng, Recent development in the treatment of oily sludge from petroleum industry: A review, Journal of Hazardous Materials 261 (2013) 470–490. https://doi.org/10.1016/j.jhazmat.2013.07.069.
[6] F. Souas, Rheological behavior of oil sludge from Algerian refinery storage tanks, Petroleum Research 7 (2022) 536–544. https://doi.org/10.1016/j.ptlrs.2022.01.002.

[7] A.Y. El Naggar, E.A. Saad, A.T. Kandil, H.O. Elmoher, Petroleum cuts as solvent extractor for oil recovery from petroleum sludge, Journal of Petroleum Technology and Alternative Fuels 1 (2010) 10–19. https://doi.org/10.5897/JPTAF.9000021.

[8] F. Nezhdbahadori, M. Abdoli, M. Baghdadi, F. Ghazban, A comparative study on the efficiency of polar and non-polar solvents in oil sludge recovery using solvent extraction, Environmental Monitoring and Assessment 190 (2018) 389. https://doi.org/10.1007/s10661-018-6748-6.

[9] M. Hassanzadeh, L. Tayebi, H. Dezfouli, Investigation of factors affecting viscosity reduction of sludge from Iranian crude oil storage tanks, Petroleum Science 15 (2018) 634–643. https://doi.org/10.1007/s12182-018-0247-9.

[10] L.J. da Silva, F.C. Alves, F.P. de França, A review of the technological solutions for the treatment of oily sludges from petroleum refineries, Waste Management & Research 30 (2012) 1016–1030. https://doi.org/10.1177/0734242X12448517.

[11] G. Hu, J. Li, H. Hou, A combination of solvent extraction and freeze–thaw for oil recovery from petroleum refinery wastewater treatment pond sludge, Journal of Hazardous Materials 313 (2016) 289–296. https://doi.org/10.1016/j.jhazmat.2016.03.073.

[12] J. Liang, L. Zhao, N. Du, H. Li, W. Hou, Solid effect in solvent extraction treatment of pre-treated oily sludge, Separation and Purification Technology 130 (2014) 28–33. https://doi.org/10.1016/j.seppur.2014.03.027.

[13] S.M. Al-Zahrani, M.D. Putra, Used lubricating oil regeneration by various solvent extraction techniques, Journal of Industrial and Engineering Chemistry 19 (2013) 536–539. https://doi.org/10.1016/j.jiec.2012.09.007.

[14] E.A. Taiwo, J.A. Otolorin, Oil recovery from petroleum sludge by solvent extraction, Petroleum Science and Technology 27 (2009) 836–844. https://doi.org/10.1080/10916460802455582.

[15] R. Ahmed, C.M. Sinnathambi, U. Eldmerdash, N-Hexane, Methyl Ethyl Ketone and Chloroform Solvents for Oil Recovery from Refinery Waste, Applied Mechanics and Materials 699 (2015) 666–671. https://doi.org/10.4028/www.scientific.net/AMM.699.666.

[16] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Networks 2 (1989) 359–366. https://doi.org/10.1016/0893-6080(89)90020-8.

[17] K. Rewatkar, D.Z. Shende, K.L. Wasewar, Reactive separation of gallic acid: Experimentation and optimization using response surface methodology and artificial neural network, Chemical and Biochemical Engineering Quarterly 31 (2017) 33–42.

[18] A.A. Mansur, M. Pannirselvam, K.A. Al-Hothaly, E.M. Adetutu, A.S. Ball, Recovery and characterization of oil from waste crude oil tank bottom sludge from Azzawiya Oil Refinery in Libya, Journal of Advanced Chemical Engineering 5 (2015) 118. https://doi.org/10.4172/2090-4568.1000118.

[19] Y. Yücel, Ö. Otuzbir, E. Yücel, Surface roughness prediction in SILAR coating process of ZnO thin films: Mathematical modelling and validation, Materials Today Communications 34 (2023) 105101. https://doi.org/10.1016/j.mtcomm.2022.105101.

[20] N. Barnier, P. Brisset, Optimisation par algorithme génétique sous contraintes, Technique et Science Informatiques 18 (1999) 1–24.