*symmetry*

MDPI

*Article*

# Using Symmetry to Model Catalyst Deactivation in an Industrial Naphtha Reforming Unit via Octane Number Prediction Using a Two-Stage Recurrent Neural Network

**Krasimir Zhekov [1,2], Darina Zhekova [3], Dicho Stratiev [1,4,*] and Svetoslav Nenov [5]**

[1] LUKOIL Neftohim Burgas, 8104 Burgas, Bulgaria; zhekov.krasimir.y@neftochim.bg
[2] Burgas State University "Prof. Dr. Assen Zlatarov", Professor Yakimov 1, 8010 Burgas, Bulgaria
[3] School of Mathematics and Natural Sciences "Academician Nikolai Obreshkov", 19 Stefan Stambolov, 8001 Burgas, Bulgaria; darinkazhekova@gmail.com
[4] Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences, Georgi Bonchev 105, 1113 Sofia, Bulgaria
[5] University of Chemical Technology and Metallurgy, Kliment Ohridski 8, 1756 Sofia, Bulgaria; s.nenov@gmail.com
* Correspondence: stratiev.dicho@neftochim.bg

**Abstract**

Catalyst activity variation during operation of a commercial catalytic naphtha reformer is a parameter that defines cycle length. The proper determination of the cycle length is vital for adequate planning in refinery operations and for optimizing the process of catalyst regeneration and turn-around scheduling. Using the symmetry between catalyst activity and the octane number of the reformate, one may model and predict catalyst deactivation over time by simulating the octane number with a two-stage recurrent neural network (RNN). In this study, data from the operation of an industrial hybrid catalytic reforming unit over a five-year period were used to develop a model based on an RNN for predicting the octane number of the reformate, as well as for simulating—based on symmetry—changes in catalyst activity and catalyst deactivation. The model allows the cycle length to be defined based on the target reformate octane number. The RNN data were separated as follows: 70% for training, 15% for validation, and 15% for testing. The first and second stages of the RNN model each had one hidden layer with seven neurons. The proposed modeling method is characterized by very high accuracy in calculating the octane number of the reformate, with a coefficient of determination ($R^2$) of 0.978 for the test data from Cycle I and 0.988 for those from Cycle II. Deactivation of the fresh catalyst was found to be much lower than that of the regenerated catalyst, and, according to the RNN model, the regenerated-catalyst cycle length is 32% shorter than that of the fresh catalyst if the same average reformate motor octane number (MON) of 88.3 is the target. However, if the same cycle length is pursued with the regenerated catalyst, then the average reformate MON should be reduced by 1.4 points.

**Keywords:** naphtha; catalytic reforming; octane number; artificial neural network; catalyst deactivation; modeling

Check for updates

Academic Editors: Cedric Karel Fonzeu Monguen, Lingnan Wu and Samuel Daniel

Received: 28 October 2026
Revised: 16 November 2026
Accepted: 19 November 2026
Published: 24 November 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

## 1. Introduction

Catalytic naphtha reforming is a key process in oil refining, contributing between 30% and 40% of the total volume of motor gasoline produced in a refinery [1]. Altering the molecular structure of the straight-run naphtha fraction, that is, the feed for the catalytic

reformer, produces high-value, high-octane reformate for the gasoline pool, hydrogen for upgrading feedstocks through hydrogen-intensive hydrotreating processes, and benzene, toluene, and xylene aromatics for petrochemical purposes. The conversion of naphthenic and paraffinic hydrocarbons into aromatic hydrocarbons determines the octane number of the reformate [2,3]. When the unit operates in gasoline production mode, the resulting reformate contains approximately 70% aromatic hydrocarbons, whereas in petrochemical mode, this content ranges between 80% and 90% [4].

The transformation of naphtha-range hydrocarbons in the reforming process is facilitated by a heterogeneous catalyst consisting of a support (typically $Al_2O_3$) onto which catalytically active metals, primarily platinum and rhenium, are deposited [5]. During operation, the catalyst loses its activity mainly due to sintering, coking, and poisoning [6–8]. Catalyst deactivation models in catalytic naphtha reforming reactions have been extensively studied [9]. Butt and Petersen have proposed a wide variety of mechanisms for explaining and modeling catalyst deactivation [10]. Quantitative correlations developed by various researchers for modeling the deactivation of reforming catalysts have been summarized by M. R. Rahimpour et al. [5]. Modeling the change in catalyst activity during operation in industrial naphtha reforming units, specifically catalyst deactivation due to coke deposition, is challenging due to the complexity of developing a rigorous deterministic model. The challenge stems from the need to account for critical process variables such as feedstock composition, variations in liquid hourly space velocity (LHSV), reactor pressure, the hydrogen-to-hydrocarbon ($H_2$/HC) molar ratio, and changes in the weight average inlet temperature (WAIT) throughout the entire operating cycle, as all of these factors influence catalyst deactivation to some extent [11].

Machine learning (ML) methods, since their introduction in 1957 [12] and especially after 2017, when the ML architecture and the overall advancement of deep learning provided robust function approximators [13,14], enabling these methods to learn the relationships between process variables in catalysis and reaction engineering research directly from the data, have become a powerful tool for modeling and optimizing chemical process performance [15–18]. In recent years, they have increasingly been recognized as a powerful data-driven approach for describing complex energy and chemical conversion systems. ML techniques such as adaptive boost, extreme gradient boost (XGBoost), categorical boost, light gradient boost (LightGBM), decision tree (DT) with bagging, random forest (RF), k-nearest neighbor, support vector machine (SVM) and artificial neural network (ANN) models have found wide application in energy-related investigations [19–22]. Sun et al. [23], investigating the advances in gasoline hydrodesulfurization (HDS) catalysts, established that ML techniques such as RF, XGBoost, LightGBM, ANNs, SVMs, and ML algorithms such as genetic algorithms (GA) or Bayesian optimization, together with regression models, have been applied to multi-objective catalyst optimization problems in refinery processes. They concluded that machine learning methods ranging from RF and SVM to deep neural networks are capable of uncovering the non-linear and multidimensional relationships between catalyst structure and performance. In the evaluation of large, heterogeneous datasets, it was found that ML models can predict catalytic behavior with greater accuracy and efficiency. Heo et al. [24], in their critical review, presented an integrated machine-learning framework that connects catalyst and processing, hydrocarbon composition, fuel properties, blending, and ASTM qualification for multi-objective optimization across the sustainable aviation fuel (SAF) production chain toward scalable 100% SAF aviation. Li et al. [25] developed a model that combines the structure-oriented lumping (SOL) technique with deep learning for the direct catalytic cracking of crude oil. A deep learning architecture (XGBoost-enhanced neural networks) brought together high-dimensional process variables (e.g., catalyst properties, reaction temperature, feedstock composition) with multi-objective outputs (light oil yield,

olefin selectivity, $CO_2$ emissions). The interpretable model is coupled with multi-objective optimization algorithms to refine critical process parameters. The selected ML techniques provided high predictive accuracy and a substantial reduction in computational cost (over sixfold). Kim et al. [26] employed seven ML models—Linear Regression, Lasso Regression, Ridge Regression, SVM, RF, Gradient Boosting, and Neural Network—to select the most appropriate multi-metal-modified ZSM-5 catalysts for bio-oil deoxygenation. An initial dataset of 146 deoxygenation reaction data samples was processed to optimize the seven ML models. The RF model demonstrated the best performance among all the models evaluated; consequently, it was used to select the most suitable catalyst combinations, resulting in the identification of the top five three-component (trimetallic) catalytic combinations. It was determined that the ML model is accurate, as the three selected catalysts proved to be the most effective for the deoxygenation of bio-oil, demonstrating excellent activity, maximum selectivity toward hydrocarbon products, and lower selectivity toward coke formation. Dizaj et al. [27] used three conventional ML models (multilayer perceptron (MLP), radial basis function (RBF), and Adaptive Neuro-Fuzzy Inference System (ANFIS)) to develop a model to predict viscosity reduction during NiO nanoparticle-assisted in situ upgrading. They compared the performance of conventional models (MLP, RBF, and ANFIS) with that of the modern tree-based ensembles, including RF and Gradient-Boosted Decision Trees (GBDT), finding that GBDT achieved the best results and demonstrated superior predictive capability compared to the other models examined. Cheenkachorn et al. [28] used XGBoost models to forecast bio-oil yield, gas composition, and char formation, demonstrating the ability of ML to accurately describe the complex processes involved in thermochemical conversion. Chotirattanachote et al. [29] developed ML models using RF, XGBoost, CatBoost, and LightGBM together with a stacking ensemble to predict total hydrocarbon yield. They found out that the best-performing models were the stacking ensemble for catalysts containing precious metals and LightGBM for catalysts containing non-precious metals. Popoola et al. [30] developed an artificial neural network (ANN) model with feed-forward architecture and back-propagation, designed to monitor the operation of a crude oil distillation column in a refinery. The authors indicate that an ANN can be effectively utilized as an ML tool for monitoring processes within a crude oil distillation unit (CDU) at an oil refinery. Umeozor [31] developed a machine learning-based (non-parametric) model by executing the extremely randomized trees regressor (ETR) method. The model was found to predict all major products with coefficients of determination exceeding 0.90, outperforming multiple linear regression and other ensemble methods. The studies reviewed above show that a wide range of machine learning techniques has been applied to oil refining and energy-related processes, with the ML method yielding the best results varying depending on the specific case. Nevertheless, among the ML techniques applied in the petroleum industry as a whole, ANNs remain the most widely used, due to their ability to describe non-linear and complex relationships that cannot be captured by linear models [32,33]. SVMs, tree-based ensembles and, for time-dependent data, recurrent architectures such as long short-term memory (LSTM) and gated recurrent unit (GRU) networks are the most common alternatives [33,34]. The same trend is observed in petroleum refining: in a comparative study of an integrated naphtha reforming and isomerization process, Gaussian process regression, ANN and SVM models predicted the research octane number of gasoline with $R^2 \geq 0.978$ and outperformed tree-based models [35].

More specifically, in catalytic reforming, ANNs have been used to predict the octane number of the reformate, the yield, and the reactor temperatures [36,37]. These ANNs include networks with recurrent layers and decay functions that account for catalyst activity [35]. However, these models primarily concern product quality, whereas the time-dependent deactivation of the catalyst has received much less attention. Al-Shathr et al. [36]

reported that a model based on a multilayer feed-forward ANN yields better results than a rigorous model of the reforming process. Sadighi et al. [37] developed a model based on a recurrent neural network (RNN), incorporating a catalyst activity decay function, to predict the octane number of the reformate and its yield. Their study demonstrates the existence of a symmetry between the octane number, reformate yield, and the decline in catalyst activity. However, they did not explore the possibility of predicting the cycle length based on this symmetry between the octane number, reformate yield, and catalyst deactivation. Stratiev et al. [38] developed models based on feed-forward deep neural networks (FFDNNs) to determine the motor octane number (MON) and the yield of reformate for two catalytic systems. In these models, catalyst activity, defined based on its relationship to the reformate octane number, which declines over the course of the operating cycle, is described using an exponential deactivation equation. To calculate the deactivation parameters (deactivation activation energy, deactivation constant, and deactivation order), the non-linear least-squares method was applied in combination with the bi-population covariance matrix adaptation evolution strategy (BIPOP-CMA-ES) algorithm. However, the cited work defines deactivation parameters only for the first catalytic system (the fresh catalyst). Deactivation parameters for the second catalytic system (the regenerated catalyst) are not provided in the study due to an insufficient amount of data. Developing RNN models that leverage the symmetry between the reformate octane number and catalyst activity, based on data for both the fresh catalyst (Cycle I) and the regenerated catalyst (Cycle II), makes it possible to determine the point at which the regenerated catalyst reaches an activity level equal to that of the fresh catalyst. To the best of our knowledge, two-stage RNN models have not yet been developed to simulate the degradation of catalyst activity over the course of an operating cycle for two catalytic systems (fresh and regenerated catalyst) in an industrial naphtha catalytic reforming unit. Consequently, this study addresses this gap by developing two-stage RNN models to determine the octane numbers of the reformate for both the first and second operating cycles with the aim of establishing the duration of the operating cycle when using a regenerated catalyst. An RNN was selected from the available machine learning methods because catalyst deactivation is a cumulative, time-dependent process. The hidden state of an RNN retains information from previous days of operation, whereas feed-forward networks, support vector machines (SVMs), and tree-based ensemble methods treat each day as an independent observation. Furthermore, a simple RNN with a single hidden layer has a small number of parameters, making it better suited to the limited size of the industrial datasets (consisting of 1,178 and 511 daily records, respectively) than deeper sequence models, such as LSTMs, GRUs, or convolutional neural networks.

The objective of this study is to develop two-stage RNN models, leveraging the symmetry between catalyst activity/deactivation and the octane number of the reformate, to determine the octane number, as well as to simulate and predict the operating cycle length for both fresh and regenerated catalysts in an industrial naphtha catalytic reforming unit.

## 2. Materials and Methods

### *2.1. Description of Studied Industrial Hybrid Reformer, Feed Characteristics and Operating Conditions*

This study, covering the period 2021–2026, focuses on a UOP-licensed hybrid catalytic naphtha reforming unit. As illustrated in Figure 1, the plant consists of three fixed-bed catalytic reactors (R1, R2, R3) and one moving-bed reactor with continuous catalyst regeneration (R4).

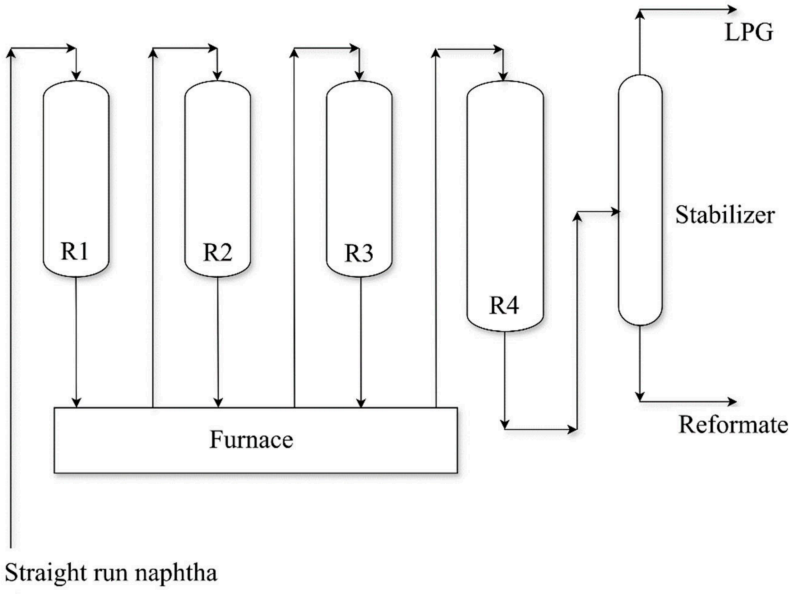



**Figure 1.** Simplified process-flow diagram of the studied hybrid industrial catalytic reforming plant.

The unit processes straight-run naphtha with a boiling point range of 100–190 °C. The feedstock undergoes hydrodesulfurization in the HDS section before entering the reactor section. The HDS step is carried out over a Co-Mo catalyst at temperatures of 310–339 °C and pressures of 26–31.2 bar. Following this pretreatment, the hydrogenated stream is fed to the platforming unit, the operating conditions of which are summarized in Table 1. Table 1 includes data about two cycles. The first one concerns the use of fresh catalyst, while the second one deals with the same catalyst used after one regeneration. Properties of both catalytic systems are summarized elsewhere [37].

**Table 1.** Operating conditions in the hybrid industrial catalytic reforming plant under investigation.

| Process Variable | Cycle I (Fresh Catalyst) Dec. 2021–Feb. 2025 (1178 Days) | | Cycle II (Regenerated Catalyst) May 2025–Oct. 2026 (511 Days) | |
|---|---|---|---|---|
| | Min | Max | Min | Max |
| Final boiling point of naphtha feed, °C | 161.9 | 190.1 | 170.9 | 193.9 |
| Naphtha feed N + 2A, wt % * | 51.6 | 69.6 | 49.6 | 72.8 |
| $P_{avg}$, bar * | 13.1 | 14.0 | 11.7 | 13.9 |
| LHSV, $h^{-1}$ * | 0.5 | 0.99 | 0.5 | 0.97 |
| $H_2$/HC, mol/mol * | 2.2 | 10.6 | 1.3 | 10.8 |
| WAIT, °C | 452 | 524.4 | 456 | 506 |
| MON | 84.3 | 89.4 | 85.2 | 89.3 |
| Days on stream | | 1178 | | 494 |
| Average reformate MON of the cycle | 88.3 | | 87.6 | |
| MON on the 511th day on stream | 89.0 | | 87.0 | |
| WAIT on the 511th day on stream, °C | 504.4 | | 506 | |
| LHSV on the 511th day on stream, $h^{-1}$ | 0.95 | | 0.95 | |

* Note: N = naphthenic hydrocarbon content, wt%; A = aromatic hydrocarbon content, wt%; $P_{avg}$ = average pressure of reactors, bar; LHSV = liquid hourly space velocity, $h^{-1}$; $H_2$/HC = hydrogen-to-hydrocarbon ratio.

Table 2 summarizes the catalyst distribution across the platforming reactors. Reactors R-1, R-2, and R-3 are charged with a bimetallic Pt–Re catalyst, whereas R-4 is loaded with a platinum catalyst.

**Table 2.** Catalyst distributions in reactors.

| Reactor | R1 | R2 | R3 | R4 |
|---|---|---|---|---|
| Catalyst weight, kg | 13,715 | 13,413 | 20,106 | 26,913 |
| Catalyst distribution, wt% | 18.5 | 18.1 | 27.1 | 36.3 |

### *2.2. Performance Calculations*

In this study, catalyst activity is calculated using a two-stage recurrent neural network (Figure 2).

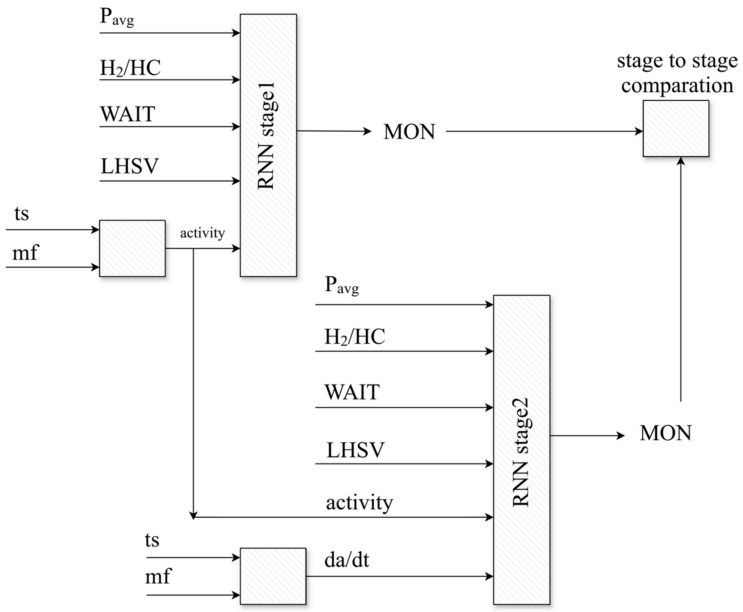



**Figure 2.** Structure of the two-stage RNN model developed for the hybrid industrial catalytic reforming plant.

In the two-stage RNN, the activity curve estimated in Stage 1 serves as an additional input variable for Stage 2, along with the derivative $\frac{d_a}{d_{t_c}}$, computed from its own ordinary differential equation (ODE), and the process variables. Each stage is based on different assumptions regarding the mechanism of catalyst activity decline. Stage 1 is governed by an exponential decay function expressed by Equation (1) and discussed in [37,39]:

$$a = e^{-\alpha \times t_c} \tag{1}$$

where $\alpha$ is the decay constant; $tc$ is the life of the catalyst, as defined in Equation (2)

$$t_c = \frac{\sum_{s=0}^{life} \left( m_f \times t_s \right)}{W_{cat}} \tag{2}$$

where $mf$ is the feed mass flow rate, kg/h; $ts$ is time on stream, h; and $Wcat$ is the catalyst inventory.

It is assumed that the loss of activity is uniform per kilogram of processed feedstock and entirely independent of the reactor temperature. Stage 2 incorporates a kinetic equation, which describes how activity varies with catalyst life. It is shown as Equation (3), and it utilizes the actual weighted average inlet temperature.

$$\frac{d_a}{d_{t_c}} = -K_d * e\left( \frac{-E_d}{R} \left( \frac{1}{WAIT} - \frac{1}{WAIT_R} \right) \right) * a^m \tag{3}$$

where $E_d$, $K_d$, $WAIT$, $WAIT_R$ and $m$ are the activation energy of catalyst deactivation, the deactivation constant, the current weight average inlet temperature, the reference $WAIT$, and the deactivation order of the catalyst, respectively. $WAIT_R$ is in fact the $WAIT$ at the start of the run of each cycle.

The parameters of Equations (1)–(3) were calculated using the best-fit reformate motor octane number (MON) obtained during the two-stage RNN model validation.

Each RNN has one hidden layer with 7 neurons. The hidden layer uses the tanh activation function (Equation (4)), while the output layer uses the linear purelin activation (Equation (5)). The data are split chronologically (in time order) into 70/15/15% for training, validation and testing.

$$tanh(n) = \frac{e^n - e^{-n}}{e^n + e^{-n}} \tag{4}$$

$$purelin(n) = n \tag{5}$$

$$n = \sum_1^{n_x} w_i * x_i + b_i \tag{6}$$

where $n$ denotes the net input of the neuron. It is obtained by multiplying each input value $x_i$ by its corresponding weight $w_i$, summing all these variables over the index range i = 1 to i = $n_x$, and adding the bias term $b$ to the resulting sum.

The cascade RNN employs two optimizers—an inner one and an outer one. The inner optimizer is based on the Levenberg–Marquardt algorithm. It trains the weights and biases of the RNN while the parameters of Equations (1)–(3) are held fixed. The outer optimizer is based on a Bayesian optimization algorithm. It searches for the parameter values in Equations (1)–(3) that minimize the mean squared error (MSE) in the calculated reformate MON. The RNN is designed for sequential data by incorporating feedback connections as shown in Figure 3.

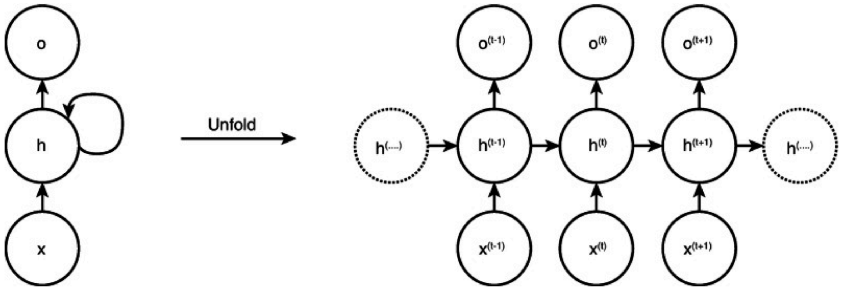



**Figure 3.** Diagram of a recurrent neural network (RNN).

Unlike feed-forward networks, RNNs retain previous information through an internal hidden state. At each time step $t$, the hidden state is:

$$h^{(t)} = f\left(h^{(t-1)}, x^{(t)}; \theta\right) \tag{7}$$

where $h^{(t)}$ is the hidden state at time step $t$, $f$ is the transition function, $h^{(t-1)}$ is the hidden state at the previous time step, $x^{(t)}$ is the input at time step $t$, and $\theta$ denotes the parameters of the RNN. When an RNN is trained to predict the future from the past, $h^{(t)}$ acts as a compressed summary of those aspects of the input sequence up to time $t$ that are relevant to the task. The recursion starts from an initial state $h^{(0)}$, and unrolling it shows that $h^{(t)}$ depends on the entire past input sequence $x^{(t)}, x^{(t-1)}, \ldots, x^{(1)}$. Importantly, the same transition function $f$ with the same parameters $\theta$ is applied at every time step. Since the model is defined as a transition between states rather than in terms of a variable-length history, its input always has the same size regardless of the sequence length, and it can generalize to sequence lengths that were not present in the training set.

In the RNN architecture considered here, the hyperbolic tangent is used as the activation function of the hidden state (Equation (8)):

$$h^{(t)} = tanh\left(a^{(t)}\right) \tag{8}$$

where $a^{(t)}$ is the argument of the activation function. Starting from the initial state $h^{(0)}$, the forward propagation for each time step from $t = 1$ to $t = \tau$ is defined by Equation (9):

$$a^{(t)} = b + W\, h^{(t-1)} + Ux^{(t)} \tag{9}$$

where $b$ is the bias vector, $W$ is the weight matrix of the recurrent (hidden-to-hidden) connections, and $U$ is the input-to-hidden weight matrix.

The output at each time step is computed from the hidden state as shown in Equation (10):

$$o^{(t)} = c + V\, h^{(t)} \tag{10}$$

where $o^{(t)}$ is the network output at time step $t$, $c$ is the bias vector of the output layer, and $V$ is the hidden-to-output weight matrix.

The parameters $\theta = \{U, W, V, b, c\}$ are learned by applying the back-propagation algorithm to the unrolled computational graph, an approach known as back-propagation through time (BPTT).

The final evaluation was done on a separate, previously unseen test set. The principal evaluation parameters were mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), the Pearson correlation coefficient (R), and the coefficient of determination ($R^2$). The formulas for their calculation are presented as Equations (11)–(15).

$$MSE = \frac{1}{n}\sum\left(MON_{exp} - MON_{calc}\right)^2 \tag{11}$$

$$RMSE = \sqrt[2]{\frac{1}{n}\sum\left(MON_{exp} - MON_{calc}\right)^2} \tag{12}$$

$$MAE = \frac{1}{n}\sum\left|MON_{exp} - MON_{calc}\right| \tag{13}$$

$$R = \frac{n\left(\sum MON_{exp}MON_{calc}\right) - \left(\sum MON_{exp}\right)\left(\sum MON_{calc}\right)}{\sqrt[2]{\left\{n\sum MON_{exp}{}^2 - \left(\sum MON_{exp}\right)^2\right\}\left\{n\sum MON_{calc}{}^2 - \left(\sum MON_{exp}MON_{calc}\right)^2\right\}}} \tag{14}$$

$$R^2 = 1 - \frac{\sum\left(MON_{exp} - MON_{calc}\right)^2}{\sum\left(MON_{exp} - MON_{average}\right)^2} \tag{15}$$

where $MON_{exp}$ = measured reformate MON; $MON_{calc}$ = reformate MON calculated by the RNN; n = number of measurements; $MON_{average}$ = average value of reformate MON.

The RNN modeling was implemented using MATLAB 2020 software.

## 3. Results and Discussion

### *3.1. Modeling the Motor Octane Number of the Reformate Using a Two-Stage RNN for Fresh Catalyst (Cycle I) and for Regenerated Catalyst (Cycle II)*

Two datasets were collected to study catalyst deactivation in an industrial hybrid naphtha reforming unit, using fresh (Cycle I) and regenerated (Cycle II) catalyst. The data for Cycle I cover 1178 days, while the data for Cycle II cover 511 days. These datasets were used to determine the optimal number of neurons in the hidden layer of two-stage RNN models designed to predict the MON of the reformate for both Cycle I and Cycle II. Over 180,000 simulations were conducted in the MATLAB 2020 environment, employing a

Bayesian optimization algorithm for both stages. Figure 4 illustrates the variation in the mean squared error (MSE) as the number of neurons changes. The data indicate that the optimal number of neurons in the two-stage RNN models for determining the MON is seven for both cycles.

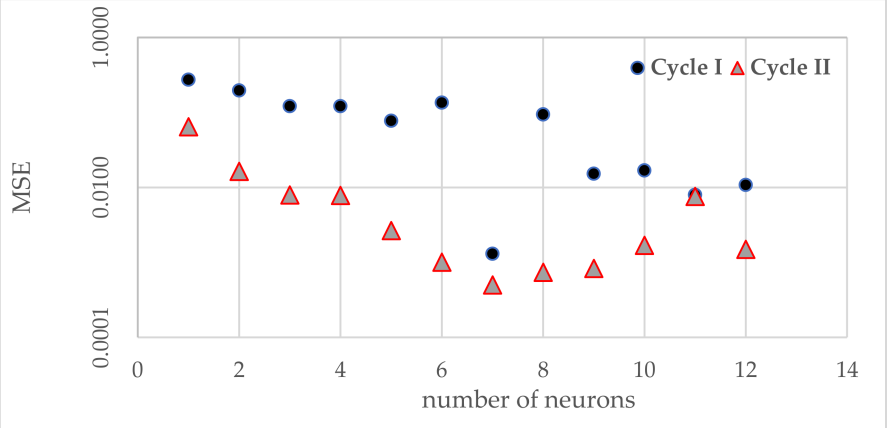



**Figure 4.** MSE of simulated MON vs. number of neurons.

The correspondence between the measured values and those calculated using the two-stage RNN model for the MON of the reformate from Cycle I during the training, validation, and testing stages is presented in Figure 5. The coefficients of determination for the three datasets are: $R^2 = 0.9316$ for training, $R^2 = 0.9976$ for validation, and $R^2 = 0.9858$ for testing, with an overall coefficient of $R^2 = 0.9928$. The statistical parameters, calculated according to Equations (11)–(15) for the two-stage RNN model of the reformate MON from Cycle I, are summarized in Table 3. These data indicate that the error in predicting the reformate MON for unseen data (from the testing stage) is lower than the reproducibility of MON measurements (0.9 units) as per the ASTM D 2700 standard [40]. Consequently, it can be concluded that the two-stage RNN model for Cycle I accurately predicts the reformate MON.

**Table 3.** Performance metrics of the two-stage RNN model for Cycle I.

| Cycle I | Data Set | MSE | MAE | RMSE | $R^2$ | R |
|---|---|---|---|---|---|---|
| two-stage RNN | Training | 0.005 | 0.051 | 0.071 | 0.932 | 0.965 |
| | Validation | 0.001 | 0.024 | 0.031 | 0.998 | 0.999 |
| | Testing | 0.025 | 0.118 | 0.159 | 0.978 | 0.989 |
| | All | 0.003 | 0.041 | 0.057 | 0.997 | 0.998 |

Figure 6 presents a parity graph comparing the measured values with those calculated using the two-stage RNN model for the MON of the Cycle II reformate, covering the training, validation, and testing stages. The coefficients of determination for the three datasets are: $R^2 = 0.984$ for training, $R^2 = 0.974$ for validation, and $R^2 = 0.988$ for testing, with an overall coefficient of $R^2 = 0.990$. The statistical parameters, calculated using Equations (11)–(15) for the two-stage RNN model of the Cycle II reformate MON, are summarized in Table 4. These data indicate that the prediction error for the reformate MON on unseen data (test data) is lower than the reproducibility of MON measurements as defined by the ASTM D 2700 standard [40]. Consequently, it can be concluded that the two-stage RNN model for Cycle II accurately predicts the MON of the reformate.

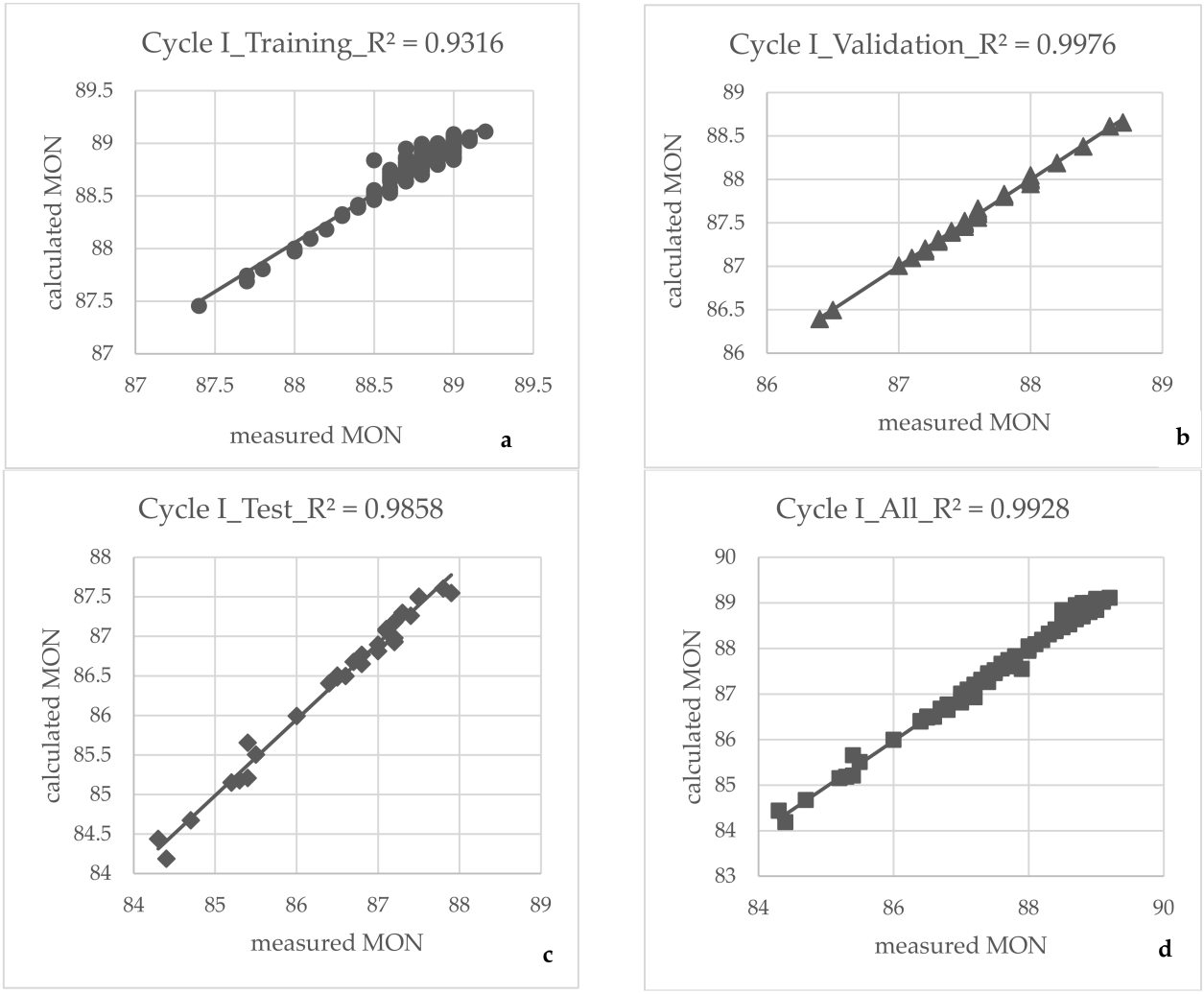



**Figure 5.** Two-stage RNN predictions versus measurements of Cycle I reformate MON: training (**a**), validation (**b**), testing (**c**), and all data (**d**).

**Table 4.** Performance metrics of the cascade RNN model for Cycle II.

| Cycle II | Data Set | MSE | MAE | RMSE | $R^2$ | R |
|---|---|---|---|---|---|---|
| two-stage RNN | Training | 0.004 | 0.033 | 0.061 | 0.984 | 0.992 |
| | Validation | 0.004 | 0.044 | 0.060 | 0.974 | 0.987 |
| | Testing | 0.001 | 0.013 | 0.014 | 0.988 | 0.994 |
| | All | 0.003 | 0.032 | 0.058 | 0.990 | 0.995 |

Comparing the results obtained in the present study regarding the prediction of the MON of the reformate—MSE = 0.025 for Cycle I (test set) and MSE = 0.01 for Cycle II (test set)—with data presented by Sadighi et al. [37] (MSE = 0.223; test set), who utilized a single-stage RNN, as well as with results reported by Stratiev et al. [38] (MSE = 0.0683; test set), who employed a deep feed-forward neural network, clearly demonstrates the advantage of the two-stage RNN concept. Given the very high accuracy in predicting the MON of the reformate using the two-stage RNN model and the symmetry between the MON and catalyst activity (and, consequently, the rate of catalyst deactivation), it can be inferred that these two catalyst characteristics could also be predicted with high accuracy.

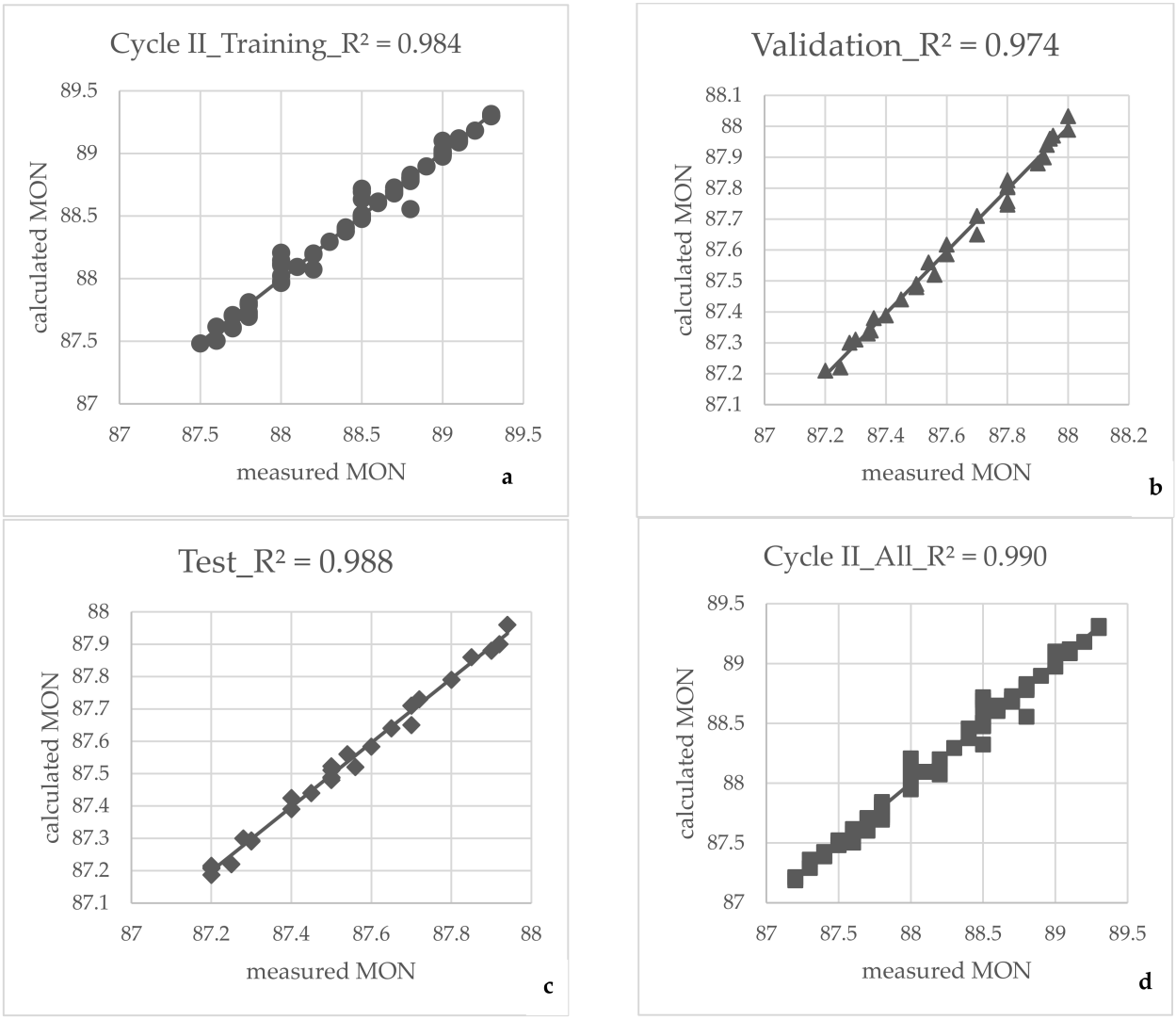



**Figure 6.** Two-stage RNN predictions versus measurements of Cycle II reformate MON: training (**a**), validation (**b**), testing (**c**), and all data (**d**).

### *3.2. Modeling of Catalyst Activity, Catalyst Deactivation, and Cycle Length Using the Two-Stage RNN MON Models for the Fresh Catalyst (Cycle I) and for the Regenerated Catalyst (Cycle II)*

The use of two-stage RNN MON models for both cycles made it possible to calculate the change in catalyst activity throughout both cycles by applying Equations (1)–(3) incorporated into these models. The variation in catalyst activity as a function of its operational duration for both cycles is presented in Figure 7. Both cycles begin with a catalyst activity equal to unity; in Cycle I, it reaches 0.84 at the 26,460 kg/kg stage, whereas Cycle II had not yet concluded at the time of this study. The data in Figure 7 clearly show a higher rate of activity decline in Cycle II (with regenerated catalyst). The deactivation constant α from Equation (1) is 6.48 for Cycle I, compared to 9.35 for Cycle II. Cycle I ends with a catalyst activity of 0.84; calculations using Equation (1) show that, if the catalyst activity in Cycle II were to reach 0.84, the cycle duration would be 376 days shorter (assuming an average naphtha feed rate of 68 t/h). The catalyst deactivation parameters *Kd*, *Ed*, and *m*, calculated using the two two-stage RNN MON models for the two cycles, are summarized in Table 5. These data show that, while the value of *Kd* for the regenerated catalyst is nearly twice as high as that for the fresh catalyst, the *Ed* and *m* values for the fresh catalyst are higher than those for the regenerated one.

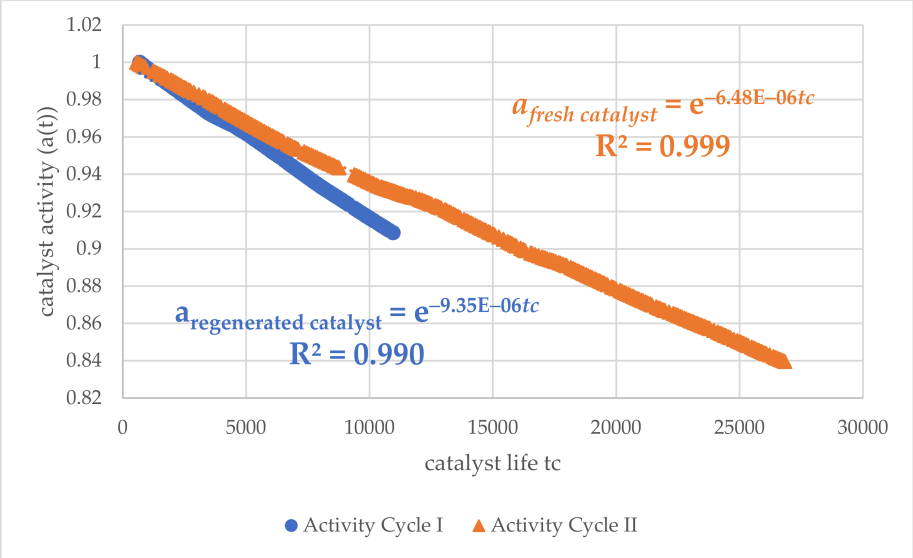



**Figure 7.** Catalyst activity in both cycles, as estimated by the 2-stage RNN MON models, versus catalyst life.

**Table 5.** Parameters of catalyst deactivation for both cycles.

| Parameter | Cycle I (2021–2025) | Cycle II (2025–2026) | Units |
|---|---|---|---|
| $K_d$ | $7.5 \times 10^{-6}$ | $13.6 \times 10^{-6}$ | $h^{-1}$ |
| $E_d$ | 161 731 | 120 400 | J $mol^{-1}$ |
| m | 6.2 | 5.3 | |

These results confirm our earlier findings, presented in [38], that the higher the value of *Ed*, the higher the order of catalyst deactivation *m*.

The cumulative effect of the different values of *Kd*, *Ed*, and *m* for the fresh and the regenerated catalysts on their deactivation rates, expressed as the respective derivatives of catalyst activity $\frac{d_{a\,fresh}}{d_{t_c}}$ and $\frac{d_{a\,regenerated}}{d_{t_c}}$, is exemplified in Figure 8.

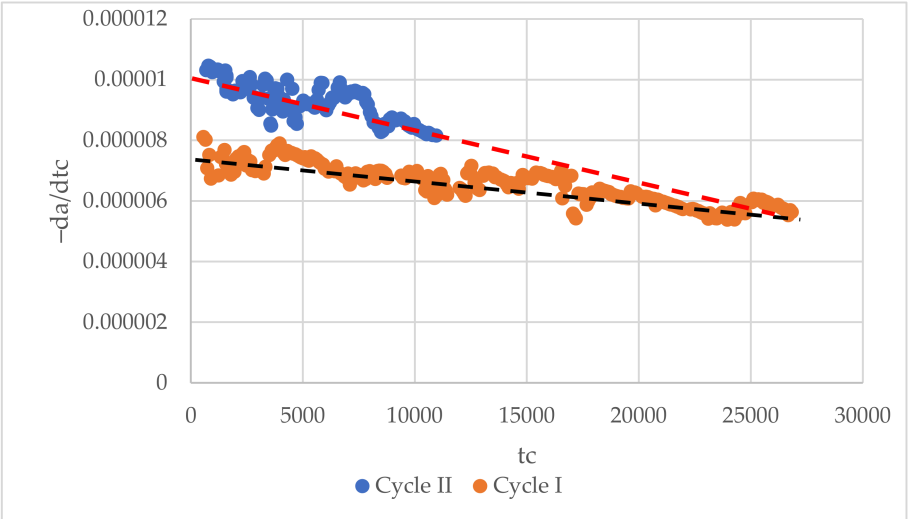



**Figure 8.** Variation in the derivative of the activity of the two catalysts—fresh and regenerated—with catalyst life. (The red dashed line represents the trend of the derivative of the activity of the regenerated catalyst, while the black dashed line represents that of the fresh catalyst.).

The data in Figure 8 (application of Equation (3)) are in full agreement with the data in Figure 7 (application of Equation (1)), showing that the regenerated catalyst exhibited a higher deactivation rate. An interesting aspect of the data in Figure 8 is the observed downward trend in the deactivation rate for both catalysts. Although the initial deactivation rate for the regenerated catalyst is higher, this decreasing trend causes it to become increasingly similar to that of the fresh catalyst over time. Extrapolating to the end of the service life (as observed during Cycle I) for the regenerated catalyst from Cycle II reveals that, by the end of Cycle II, the two catalysts have identical deactivation rates. One possible explanation for this is that, as the service life progresses, regardless of whether the catalyst is fresh or regenerated, an increasing number of active sites become blocked by coke deposited on the surface. Consequently, fewer accessible active sites remain that could be subject to further coking and blockage. It appears that, by the end of the cycle, both the fresh and the regenerated catalysts exhibit the same deactivation rate.

Given that the catalyst deactivation rate increases exponentially with a rise in the WAIT, as indicated in [37], calculations using Equation (10) from the same source show that increasing the WAIT from 506 °C (per Table 1) to the maximum of 524.4 °C (see Table 1) would take 17 months if the reformate MON were maintained at a level of 87.0. Considering that the average reformate MON on day 511 of Cycle II is 87.6, it can be readily calculated that, by day 1178 (the day on which Cycle I ended), the average reformate MON for the entire cycle will be 86.9. A comparison of the average reformate MON of 86.9 for Cycle II with the value of 88.3 for Cycle I (see data in Table 1) indicates that, due to the more rapid deactivation of the regenerated catalyst, the average reformate MON will be lower by less than 1.4 points.

## 4. Conclusions

This study demonstrates that applying symmetry when modeling catalytic activity and the catalyst deactivation rate by utilizing a two-stage RNN model to determine the MON of the reformate enables the estimation of the cycle length. It was found that the two-stage RNN model for determining the MON is more accurate than both the single-stage RNN model described in the scientific literature and a deep feed-forward neural network. This suggests that the predicted values for catalytic activity and the deactivation rate, derived from the two-stage RNN model, should also exhibit a very high degree of accuracy. The results of this research make it possible to predict the cycle length of a regenerated catalyst, as well as the implications of its shorter service life compared to that of a fresh catalyst. This includes assessing the economic impact of regenerating the catalyst earlier than planned versus extending the operating cycle, a scenario that results in a lower MON for the reformate and necessitates the additional purchase of expensive high-octane components, such as methyl tert-butyl ether (MTBE). These findings offer practical utility: an oil refinery can proactively plan catalyst regeneration and maintenance activities. Furthermore, it is possible to evaluate the trade-off between the octane number of the reformate and the length of the operating cycle. The proposed method could also be applied to other oil refining units, such as various hydroprocessing units.

**Author Contributions:** Conceptualization, D.S. and K.Z.; methodology, K.Z.; software, D.Z.; validation, S.N., K.Z. and D.S.; formal analysis, K.Z.; investigation, D.Z.; resources, D.S.; data curation, S.N.; writing—original draft preparation, D.S. and K.Z.; writing—review and editing, D.S. and K.Z.; visualization, S.N.; supervision, D.S.; project administration, K.Z.; funding acquisition, K.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

**Acknowledgments:** The authors acknowledge the support given by their colleagues Dobromir Yordanov and Sotir Sotirov during the preparation of this manuscript. During the preparation of this manuscript, the authors used Free AI writer, https://quillbot.com/ai-writing-tools/ai-writer (access on 18 November 2026), to polish the English in the manuscript. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

**Conflicts of Interest:** The authors Dicho Stratiev and Krasimir Zhekov were employed by the company LUKOIL Neftohim Burgas. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| ANFIS | Adaptive neuro-fuzzy inference system |
| ANN | Artificial neural network |
| ASTM | American Society for Testing and Materials |
| BIPOP-CMA-ES | Bi-population covariance matrix adaptation evolution strategy |
| BPTT | Back-propagation through time |
| FFDNN | Feed-forward deep neural network |
| GBDT | Gradient-boosted decision trees |
| GRU | Gated recurrent unit |
| $H_2$/HC | Hydrogen-to-hydrocarbon molar ratio |
| HDS | Hydrodesulfurization |
| LHSV | Liquid hourly space velocity |
| LightGBM | Light gradient boosting machine |
| LSTM | Long short-term memory |
| MAE | Mean absolute error |
| ML | Machine learning |
| MLP | Multilayer perceptron |
| MON | Motor octane number |
| MSE | Mean squared error |
| ODE | Ordinary differential equation |
| R | Pearson correlation coefficient |
| $R^2$ | Coefficient of determination |
| RBF | Radial basis function |
| RF | Random forest |
| RMSE | Root mean squared error |
| RNN | Recurrent neural network |
| SAF | Sustainable aviation fuel |
| SVM | Support vector machine |
| UOP | Universal Oil Products |
| WAIT | Weight average inlet temperature |
| XGBoost | Extreme gradient boosting |

## References

1. Oyekan, S. *Catalytic Naphtha Reforming Process*; CRC Press: Boca Raton, FL, USA, 2018. [CrossRef]
2. Al-Shafei, E.N.; Albahar, M.Z.; Aljishi, M.F.; Akah, A.; Aljishi, A.N.; Alasseel, A. Catalytic conversion of heavy naphtha to reformate over the phosphorus-ZSM-5 catalyst at a lower reforming temperature. *RSC Adv.* **2022**, *12*, 25465–25477. [CrossRef]
3. Moljord, K.; Hellenes, H.G.; Hoff, A.; Tanem, I.; Grande, K.; Holmen, A. Effect of reaction pressure on octane number and reformate and hydrogen yields in catalytic reforming. *Ind. Eng. Chem. Res.* **1996**, *35*, 99–105. [CrossRef]
4. Fahim, M.A.; Alsahhaf, T.A.; Elkilani, A. Chapter 5–Catalytic Reforming and Isomerization. In *Fundamentals of Petroleum Refining*; Elsevier: Amsterdam, The Netherlands, 2010; pp. 95–122. [CrossRef]
5. Rahimpour, M.R.; Jafari, M.; Iranshahi, D. Progress in catalytic naphtha reforming process: A review. *Appl. Energy* **2013**, *109*, 79–93. [CrossRef]
6. Fogler, H.S. *Elements of Chemical Reaction Engineering*, 5th ed.; Prentice Hall: Boston, MA, USA, 2016.

7. Jess, A.; Hein, O.; Kern, C. Deactivation and decoking of a naphtha reforming catalyst. In *Studies in Surface Science and Catalysis*; Delmon, B., Froment, G.F., Eds.; Elsevier: Amsterdam, The Netherlands, 1999; Volume 126; pp. 81–88. [CrossRef]
8. Yusuf, A.Z.; Aderemi, B.O.; Patel, R.; Mujtaba, I.M. Study of Industrial Naphtha Catalytic Reforming Reactions via Modelling and Simulation. *Processes* **2019**, *7*, 192. [CrossRef]
9. Ali, S.A.; Alshareef, A.H.; Theravalappil, R.; Alasiri, H.S.; Hossain, M.M. Molecular Kinetic Modeling of Catalytic Naphtha Reforming: A Review of Complexities and Solutions. *Catal. Rev.* **2023**, *65*, 1358–1411. [CrossRef]
10. Butt, J.B.; Petersen, E.E. *Activation, Deactivation, and Poisoning of Catalysts*; Academic Press: San Diego, CA, USA, 1988.
11. Mohadesi, M.; Mousavi, H.S. Kinetic Investigation of Catalyst Deactivation in Catalytic Reforming of Naphtha. *Chem. Prod. Process Model.* **2012**, *7*, 9. [CrossRef]
12. Fradkov, A.L. Early History of Machine Learning. *IFAC-PapersOnLine* **2020**, *53*, 1385–1390. [CrossRef]
13. Larraz, R.; Corma, A. From Atoms to Reactors: A Critical Review of a Decade of Machine Learning in Heterogeneous Catalysis (2017–2026). *Angew. Chem. Int. Ed.* **2026**, e4288841. [CrossRef]
14. Gao, Y.; Hong, J.; Wang, P.; Guo, C.; Yan, H.; Wang, Z.; Ma, J.; Zhang, Q.; Tang, D.; Bo Jiang, B.Q. Interpretable Descriptor Enabling Prediction of Propylene Yield for Propane Dehydrogenation Catalysts. *Adv. Energy Mater.* **2026**, *16*, e70893. [CrossRef]
15. Günay, M.E.; Yıldırım, R. Machine Learning for Catalytic Reaction Systems: A Framework for Complex Chemical Processes. *ACS Eng. Au* **2026**, *6*, 48–67. [CrossRef]
16. Bratovčić, A.; Tomašić, V. Artificial Intelligence- and Machine Learning-Driven Strategies for Catalyst Design and Sustainable Chemical Processes. *Processes* **2026**, *14*, 1866. [CrossRef]
17. Li, N.; Xu, J.; Gao, W.; Wan, S.; Xu, X.; Yan, B.; Wang, J.; Chen, G. Machine learning-enabled advanced oxidation processes: Catalytic active site engineering, oxidation pathway regulation, and reaction condition optimization. *Chem. Eng. J.* **2026**, *528*, 172413. [CrossRef]
18. Wang, S.; Chau, H.; Kristy, S.; Thompson, B.A.; Malizia, J.P.; Fushimi, R. Machine Learning-Assisted Recovery of Delicate Kinetic Information from Transient Reactor Experiments. *ACS Eng. Au* **2025**, *5*, 298–310. [CrossRef]
19. Mafat, I.H.; Sharma, S.K.; Surya, D.V.; Rao, C.S.; Maity, U.; Barupal, A.; Jasra, R. Development of Machine Learning Model for the Prediction of Selectivity to Light Olefins from Catalytic Cracking of Hydrocarbons. *Fuel* **2025**, *381*, 133682. [CrossRef]
20. Acosta-López, J.G.; de Lasa, H. Artificial Intelligence for Hybrid Modeling in Fluid Catalytic Cracking (FCC). *Processes* **2024**, *12*, 61. [CrossRef]
21. Mishra, P.; Yadav, A. CFD-Deep Neural Network (DNN) Model of a Slurry Phase Reactor for Vacuum Residue Hydrocracking. *Int. J. Hydrogen Energy* **2025**, *164*, 150726. [CrossRef]
22. Pizoń, Z.; Kimijima, S.; Brus, G. Bridging Equilibrium and Kinetics Prediction with a Data-Weighted Neural Network Model of Methane Steam Reforming. *Int. J. Hydrogen Energy* **2025**, *175*, 151367. [CrossRef]
23. Sun, H.; Chen, C.; Zhang, R.; Li, Y.; Ge, S.; Cui, P. Advances in Gasoline Hydrodesulfurization Catalysts: The Role of Structure–Activity Relationships and Machine Learning Approaches. *ACS Omega* **2025**, *10*, 31262–31273. [CrossRef]
24. Heo, S.; Lim, S.Y.; Paulson, J.A.; Huber, G.W.; Kim, Y.T.; Lee, J. Machine learning-enabled bio-oil co-processing for sustainable aviation fuel: Catalyst-to-refinery integration for blend expansion. *Green Chem.* **2026**, *28*, 16122–16190. [CrossRef]
25. Li, Z.; Qian, X.; Liu, T.; Wang, Y.; Huang, X.; Duan, H.; Zhang, Z.; Sun, Z.; Yan, H.; Chen, X.; et al. Enhancing engineering strategies for direct catalytic cracking of crude oil via molecular-level kinetic reaction networks: A machine learning-enabled multi-objective optimization framework. *Fuel* **2026**, *413*, 138273. [CrossRef]
26. Kim, G.; Liu, A.; Jarvis, J.; Bonatti, E.; Li, Z.; Song, H. Machine learning assisted bio-oil deoxygenation under methane environment: The model compound study on guaiacol. *Catal. Today* **2025**, *459*, 115428. [CrossRef]
27. Ghodrati Dizaj, A.; Namdar, H.; Jafari, A. Comparative analysis of artificial intelligence models for predicting oil viscosity in the in-situ catalytic oil upgrading process: A study of influential parameters. *Sci. Rep.* **2026**, *16*, 31744. [CrossRef]
28. Cheenkachorn, K.; Prapainainar, C.; Wijakmatee, T. Machine learning-driven modeling of biomass pyrolysis product distribution through thermal parameter sensitivity. *Renew. Energy* **2025**, *248*, 123108. [CrossRef]
29. Chotirattanachote, A.; Thiensuwan, N.; Ngamcharussrivichai, C.; Yokoi, T.; Sukpancharoen, S. Machine learning-guided catalyst selection for bio-oil hydrodeoxygenation to maximize hydrocarbon yields toward renewable naphtha and sustainable aviation fuel. *Fuel* **2026**, *426*, 139511. [CrossRef]
30. Popoola, L.T.; Al-Saidi, A.; Asmara, Y.P.; Taura, U.; Obende, B.A.; Osibuamhe, M.O. Artificial Neural Network Model for Process Monitoring of Crude Oil Distillation Column in a Petroleum Refinery. *Concurr. Eng. Res. Appl.* **2025**, *33*, 79–92. [CrossRef]
31. Umeozor, E. Application of Machine Learning Models to Oil Refinery Programming. *Processes* **2026**, *14*, 1072. [CrossRef]
32. Sircar, A.; Yadav, K.; Rayavarapu, K.; Bist, N.; Oza, H. Application of machine learning and artificial intelligence in oil and gas industry. *Pet. Res.* **2021**, *6*, 379–391, Erratum in *Pet. Res.* **2023**, *8*, 283–284. https://doi.org/10.1016/j.ptlrs.2023.05.007. [CrossRef]
33. Salem, A.M.; Yakoot, M.S.; Mahmoud, O. Addressing Diverse Petroleum Industry Problems Using Machine Learning Techniques: Literary Methodology—Spotlight on Predicting Well Integrity Failures. *ACS Omega* **2022**, *7*, 2504–2519. [CrossRef]

34. Mai-Cao, L.; Truong-Khac, H. A Comparative Study on Different Machine Learning Algorithms for Petroleum Production Forecasting. *Improv. Oil Gas Recovery* **2022**, *6*, 1–8. [CrossRef]
35. Saghir, H.; Ahmad, I.; Kano, M.; Caliskan, H.; Hong, H. Prediction and optimisation of gasoline quality in petroleum refining: The use of machine learning model as a surrogate in optimisation framework. *CAAI Trans. Intell. Technol.* **2024**, *9*, 1185–1198. [CrossRef]
36. Al-Shathr, A.; Shakor, Z.M.; Majdi, H.S.; AbdulRazak, A.A.; Albayati, T.M. Comparison between artificial neural network and rigorous mathematical model in simulation of industrial heavy naphtha reforming process. *Catalysts* **2021**, *11*, 1034. [CrossRef]
37. Sadighi, S.; Seif Mohaddecy, R.; Norouzian, A. Optimizing an industrial scale naphtha catalytic reforming plant using a hybrid artificial neural network and genetic algorithm technique. *Bull. Chem. React. Eng. Catal.* **2015**, *10*, 210–220. [CrossRef]
38. Stratiev, D.; Petrov, I.; Dinkov, R.; Nedelchev, A.; Shishkova, I.; Shishkov, K.; Zhekov, K.; Ivanov, M.; Nenov, S.; Nikolova, R. Impact of Feed Composition and Operating Conditions on Product Yield and Quality, and Cycle Length of an Industrial Hybrid Naphtha Catalytic Reformer. *ChemEngineering* **2026**, *10*, 128. [CrossRef]
39. Jia, B.; Wang, S.; Yang, J.; Huang, Z. Multi-stage kinetic analysis and product distribution prediction for oil shale pyrolysis. *Fuel* **2026**, *420*, 138994. [CrossRef]
40. *ASTM D2700-22*; Standard Test Method for Motor Octane Number of Spark-Ignition Engine Fuel. ASTM International: West Conshohocken, PA, USA, 2022.


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.